DOCUMENT RESUME

ED 314 059 IR 052 988

AUTHOR Fidel, Raya
TITLE Extracting Knowledge for Intermediary Expert Systems: The Selection of Search Keys. Final Report.
INSTITUTION Washington Univ., Seattle. Graduate School of Library and Information Science.
SPONS AGENCY National Science Foundation. Washington, D.C. Div. of Information Science and Technology.
PUB DATE Jun 88
GRANT IST-85-09719
NOTE 124p.
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Bibliographic Databases; Case Studies; Expert Systems; *Models; *Online Searching; Scientific and Technical Information; *Search Strategies; *Subject Index Terms; Thesauri
IDENTIFIERS *Free Text Searching

ABSTRACT

This study investigated online searching behavior manifested by 39 experienced professional searchers performing their regular, job-related searches in order to uncover the rules they use for the selection of search keys, and to represent these rules in a formal model that could be used in the construction of intermediary expert systems. The case study method with controlled comparison was used, and data analyses were based on two existing models: the Selection Routine, a decision tree presenting the rules used to select search keys by eight searchers in a previous study, and a list of moves, or modifications, in search strategies that was based on observations of the searching behavior of the same eight subjects. Two types of moves were identified: operational moves that preserve the meaning of a request, and conceptual moves that change the meaning of a request. Within each type, the moves are presented in three groups: precision moves; recall moves; and moves to increase precision and recall. Data analysis involved measuring the frequency with which each type of search key was selected, each move was selected, and a reason was cited to explain the selection of a search key. The statistical associations among 11 variables were also examined for each search: (1) number of search keys selected; (2) percentage of free-text terms used; (3) frequency with which a thesaurus was not consulted; (4) number of databases used; (5) number of moves made; (6) percentage of operational moves made; (7) number of precision moves made; (8) number of recall moves made; (9) percentage of recall moves made; (10) the searcher's subject area specialty; and (11) the searcher's work environment. It was concluded that searching behavior is lawful and follows certain patterns; current systems cannot provide satisfactory recall and some are even an impediment to useful searching practices; and the requests presented by users are the least predictable of the factors that affect searching behavior. (6 tables, 3 figures, and 42 references) (SD)

ED314059

EXTRACTING KNOWLEDGE FOR INTERMEDIARY EXPERT SYSTEMS:

THE SELECTION OF SEARCH KEYS

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Final Report

for

National Science Foundation Grant IST 85-09719

Raya Fidel

Graduate School of Library and Information Science

University of Washington

Seattle, WA 98195

June 1988

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Raya Fidel

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IR052988

BEST COPY AVAILABLE

TABLE OF CONTENTS

Executive Summary. iv

Acknowledgments. ix

List of Tables. x

List of Figures. xi

1. INTRODUCTION. 1

  1.1 Problem Definition. 2

    1.1.1 Controlled Vocabulary and Free-text Keys. 3
    1.1.2 Intermediary Expert Systems. 4

  1.2 The Objectives of the Study. 8

    1.2.1 Refinement and Validation of the Selection Routine. 8
    1.2.2 The Effect of Searching Behavior on Search-Key Selection. 9
    1.2.3 The Applicability of the Case Study Method. . . . 9

2. THE METHOD. 10

  2.1 Data Collection. 11

    2.1.1 The Procedure. 11
      (a) Contacting the Searcher. 11
      (b) Recording Sessions. 12
      (c) Transcribing Recorded Material. 13
      (d) Analyzing Search Protocols and Verbalizations. .13
      (e) Recording Data on Forms. 14
      (f) Participating in Weekly Meetings. 15
      (g) Interviewing the Searcher. 16
    2.1.2 Adjustments. 16
      (a) The Selection of Cases. 16
      (b) Retrospective Observations. 17
      (c) Initial Contact with Searchers. 18
      (d) Co-observation. 18

  2.2 The Selection of Searchers. 19

  2.3 Data Analysis. 21

2.4 Advantages and Limitations of the Method. 24

3. ONLINE SEARCHING BEHAVIOR. 25

3.1 The Selection Routine. 26

3.1.1 A Term is a Common Term. 29
3.1.2 A Single-Meaning Term That is Mapped to a Descriptor. 30
3.1.3 A Single-Meaning Term That is Not Mapped to a Descriptor. 33
3.1.4 It is Not Known If a Term is Mapped to a Descriptor. 34

3.2 Searchers' Selection of Search Keys. 36

3.2.1 Frequency of Search Key Selection. 36
3.2.2 Frequency of Option Selection. 36
3.2.3 Frequency of Reasons for Option Selection. 39
(a) Request-Related Reasons. 42
(b) Database-Related Reasons. 43
(c) Searcher-Related Reasons. 44
3.2.4 Frequency of Search-Key Selection for Databases. . 45

3.3 Searchers' Selection of Moves. 48

3.3.1 The Frequency of Moves Selection. 48

4. FACTORS AFFECTING THE SELECTION OF SEARCH KEYS. 52

4.1 The Number of Search Keys. 54

4.1.1 The Number of Moves. 54
4.1.2 Environment. 54
4.1.3 The Number of Databases. 55
4.1.4 Search-Keys Ratio. 55
4.1.5 Moves Ratio. 56
4.1.6 Recall Tendency. 56
4.1.7 Subject Area. 57

4.2 Search-Keys Ratio. .58

4.2.1 The Number of Databases. 58
4.2.2 Moves Ratio. 58
4.2.3 Subject Area. 59
4.2.4 The Individual Databases. 59
4.2.5 Environment. 60
4.2.6 The Number of Moves. 60

4.3 Thesaurus Look-Ups. 62

4.3.1 The Number of Databases. 62
4.3.2 Moves Ratio. 63
4.3.3 Subject Area. 63
4.3.4 The Number of Search Keys. 64

4.3.5 The Number of Moves. 65
4.3.6 Search-Keys Ratio. 65
4.3.7 Unrelated Variables. 65

4.4 The Number of Databases. 66

4.4.1 Subject Area. 66
4.4.2 Moves. 66
4.4.3 Moves Ratio. 67
4.4.4 Environment. 67

4.5 The Number of Moves. 68

4.6 The Moves Ratio. 69

4.6.1 Subject Area. 69
4.6.2 Moves. 69
4.6.3 Environment. 70

4.7 Recall Tendency. 71

5. SUMMARY AND CONCLUSIONS. 73

5.1 The Selection Routine. 75

5.1.1 The Selection of Options. 75
5.1.2 Thesauri Quality and Availability. 76
5.1.3 The Concern with Recall. 77

5.2 Factors Affecting Searching Behavior. 79

5.2.1 The "Free-Text" Searcher. 79
5.2.2 Factors Typical of Searching Behavior. 80
5.2.3 The Effect of Requests on Searching Behavior. 81
5.2.4 The Effect of Design Factors. 83

5.3 The Case Study Method. 85

5.4 Implications for Future Research. 86

6. REFERENCES. 87

APPENDICES

Executive Summary

Final Report on NSF Grant ITS 85-09719

EXTRACTING KNOWLEDGE FOR INTERMEDIARY EXPERT SYSTEMS:
THE SELECTION OF SEARCH KEYS

Raya Fidel

The purpose of this study was to uncover the rules used by online searchers for the selection of search keys, whether free-text terms or descriptors, and to represent these rules in a formal model that could be used in the construction of intermediary expert systems.

The case study method with controlled comparison was used to analyze the data which were collected through observation of online searchers performing their regular, job-related searches. The study's participants were experienced searchers who were selected from a wide spectrum of subject specialties and from various settings.

Data analysis was based on two existing models. The first was the original Selection Routine, a decision tree that presented the rules used to select search keys by eight searchers in a previous study. The second was a list of moves--modifications in search strategies--based on observing the searching behavior of the same eight searchers. The list is divided into two types of moves: operational moves which keep the meaning of a request unchanged; and conceptual moves which change the meaning of a request. Within each type, the moves are presented in three groups: precision moves; recall moves; and moves to increase both precision and recall.

This study included 39 searchers whose searching behavior was analyzed in order to expand both models. These searchers were also asked to explain their reasons for each selection of a search key.

Data analysis involved measuring the frequency with which: (1) each type of search key was selected; (2) each move was selected; and (3) a reason was cited to explain the selection of a search key. Further, the statistical associations among eleven variables were examined. These variables are: (1) the number of search keys selected for a search; (2) search-keys ratio (the percentage of free-text terms compared to the total number of search keys); (3) thesaurus look-ups (the frequency with which a thesaurus was not consulted); (4) the number of databases used per search; (5) the number of moves made in a search; (6) moves ratio (the percentage of operational moves compared to the total number of moves); (7) the number of precision moves made in a search; (8) the number of recall moves made in a search; (9) recall tendency (the percentage of recall moves compared to the total number of moves); (10) the subject area in which a searcher specializes; (11) the environment

in which a searcher works.

The statistical analyses included data from the 39 study searchers as well as from the eight original searchers (a total of 47), and were performed on two levels: the search level, in which each search was considered as an instance (281 instances); and the person level, in which data from all the searches performed by one person were aggregated to represent a single instance (47 instances).

The study had three explicit objectives: (1) to refine and validate the Selection Routine; (2) to explore the effect of searching behavior on search-key selection; and (3) to test the applicability of the case study method to the extraction of knowledge from multiple experts.

1. The Selection Routine

The Selection Routine was modified by the analysis of searching behavior. Although the modified Selection Routine is not complete (because a few infrequent options are still unexplained), it can be used to develop the set of rules for a rule-based intermediary expert system.

The selection of options. The frequency with which search keys were selected and the reasons cited by the searchers for their selection of these keys revealed that:

--Searchers did not display a general preference for one type of search key: When they had a choice, they selected descriptors and free-text terms in the same frequency.

--About 70% of the time, searchers selected the most straightforward options: If a term were exactly mapped to a descriptor, they entered a descriptor; if it could not be mapped, or they did not consult a thesaurus, they entered free-text terms.

--When searchers had options in the selection of search keys, their choice was most frequently (48%) determined by the databases they were searching, less frequently (32%) by the request they were searching, and least frequently (20%) by their habitual searching behavior.

Thesauri quality and availability. All 47 searchers relied heavily on thesauri: They consulted a thesaurus for 80% of the search keys they selected, and when they did not consult a thesaurus, it was often because a thesaurus was unavailable. Further, the quality of thesauri and indexing, and their availability, greatly affect the selection of search keys:

--When the characteristics of a database were cited as a reason for selection a particular search key, 25% of the reasons given referred to the lack of a needed descriptor, 19% to thesaurus unavailability, and 18% pinpointed the poor quality of descriptors and indexing.

--Distrust of descriptors and/or indexing explained 16% of the instances in which searchers entered terms *without* consulting a thesaurus.

Therefore, the quality and availability of thesauri are critical factors in the selection of search keys, and better quality in thesauri and indexing, as well as greater availability of these tools, are badly needed.

The concern with recall. The searchers who participated in the study were heavily occupied with attempts to increase recall:

--If searchers did not initially choose a straightforward option in the selection of search keys, more than half the time they chose an option that would enhance recall.

--The most frequent reason for the selection of a search key that referred to the request was the need to enhance recall (35%).

--The number of moves to increase recall was almost double the number of moves to increase precision.

These results snow that with the current bibliographic databases it is difficult to achieve recall scores that are satisfactory to searchers.

2. The Effect of Searching Behavior

The results of the study revealed the factors that affect the selection of search keys.

The "free-text" searcher. The statistical tests show that a profile can now be constructed of the searchers who use free-text terms more often than other searchers. They are likely to have these characteristics:

--be an operationalist searcher (that is, make operational moves more frequently than conceptual ones),

--be a searcher in the sciences,

--if, as a science searcher, they usually answer practical requests, they will use even more free-text terms,

--need to search several databases for each request, and have developed a habit of entering terms without consulting a thesaurus.

Contrary to common notions, searchers who prefer to enter free-text terms do not enter more search keys than those who prefer descriptors, nor are they more interactive than their counterparts. That is, searchers who prefer to use free-text terms most often enter these terms as if they were descriptors, without exercising terminological control in searching.

Factors typical of searching behavior. Statistical analyses discovered that some searchers are routinely more interactive than others: they make more moves, they enter more search keys, and they use more databases than their peers who are less interactive.

In addition, the searching style of a searcher, whether operationalist or conceptualist, also affects the selection of search keys and other aspects of searching behavior. Operationalist searchers:

--use free-text terms more frequently,

--are more likely to avoid consulting a thesaurus,

--are more likely to answer science or general questions, and

--are more likely to make precision moves than conceptualist searchers.

The effect of requests. The nature of each individual request affects the selection of search keys: 32% of the times they explained their search-keys selection, the study's searchers referred to requirements of the request. Since 48% of the reasons related to constraints of the databases, it is plausible to assume that with improved flexibility in the structure of thesauri, and in availability of searching tools, searchers will give request characteristics higher priority.

Further, requests that present terminological difficulties (as reflected by the use of a relatively large number of search keys), and those that are difficult to search (as measured by the number of moves that are made to answer them) lead searchers to enter search keys without consulting a thesaurus. Thus, requests with terminological difficulties require more interaction than other requests, an interaction during which searchers add search keys without consulting a thesaurus.

In addition, the subject area and whether a request is practical or theoretical also determine the selection of search keys:

--Science requests are searched with free-text terms more frequently than requests in other subject areas.

--Science and general requests are more likely to be searched by operational moves than by conceptual ones.

--Practical requests are searched with free-text terms more frequently than theoretical ones.

--Theoretical requests require higher recall than practical ones. The results of the study, however, could not substantiate the common belief that high-recall requests require an increased number of search keys.

The effect of design factors. Among the variables examined in this study, only "the number of databases used in a search" relates to design factors because it is determined by the distribution of information among the databases, and it is, therefore, a given with searchers.

Study results show that this variable correlates with almost all other variables, indicating that the number of databases a searcher has to use for a search has crucial effect on the selection of search keys and on other aspects of online searching behavior. Most notable is the discovery that having to search several databases for a request induces the use of free-text terms and entering these terms without consulting a thesaurus. Therefore, having to search a variety of databases for one request is a limiting factor.

Though it is generally believed that the availability of a large number of databases enhances online searching because searchers have more choice, it is not clear how much freedom is actually introduced by this multitude--each database is somewhat different from the others and searchers often feel they must use every database possible to be comprehensive. On the other hand, the use of a variety of databases limits their options in the selection of search keys, because preparing a different search strategy for each individual databases is unrealistically time consuming.

The uncoordinated growth of databases is an impediment to online searching. Standardization and coordination are required if users are to be able to fully exploit the capabilities of online systems.

3. The Case Study Method

The applicability of the case study method to the extraction of knowledge from a number of experts is proven by the success of the method to create formal models of the selection of search keys and of moves.

Further, the use of the method in this study led to two conclusions: (1) the method of controlled comparison is useful for resolving conflicting evidence; and (2) observation and analysis of a relatively small number of searchers but with a variety of backgrounds, is sufficient for the creation of a model that describes their searching behavior in formal terms.

4. Implications for Future Research

The most timely conclusion of this study is the proof it provides that searching behavior is not completely determined by individual idiosyncrasies, but rather that searching behavior is lawful and follows certain patterns.

More specific findings of the study point to the need to investigate the impediments to achieving satisfactory recall, the characteristics of requests that affect the search process, and the ways in which current thesauri and indexing practices could be improved.

ACKNOWLEDGMENTS

The searchers who participated in this study were extremely cooperative spirits. Despite their heavy schedules and the time pressure under which they perform their jobs, they worked willingly with the study team, and even adjusted their schedules at times to accommodate the team's needs. Their cooperation made this study possible.

Members of the study team were Nancy Phelps, Michael Crandall, Cynthia Altick Cunningham, and Kathleen McCrory. They rapidly became professional observers and critical analysts, and are responsible for the high quality of the data that have been collected. Their contribution to the study is substantial and highly appreciated.

In addition, I would like to thank: my colleague Terrence Brooks, for his help with the statistical analyses; the National Science Foundation, Division of Information Science and Technology, Program in Information Technology, for financial support; and, in particular, Dr. Harold Bamford, project officer the Program in Information Technology, for his constructive comments on the plan of the study.

List of Tables

Table 1.--A List of Options and the Associated Conditions. 28

Table 2. Frequency of Options and Reasons. 37

Table 3. Search-Key Selection for Databases. 46

Table 4. Moves in Online Searching. 49

Table 5. Frequency of Move Selection. 50

Table 6. Summary of the Factors Affecting Searching Behavior. . . .72

List of Figures

Figure 1. The Selection Routine--a Decision-Tree Display. 27

Figure 2. The Selection Routine--a Network Display. 40

Figure 3. Frequent Options in the Selection Routine --a Network Display. 41

1. INTRODUCTION

Online searching behavior has attracted much attention among researchers because of the current discrepancy between the level of technological developments as compared to theoretical advancements. New and increasingly sophisticated technology is being developed and put to use at an ever-growing rate, but the scientific understanding of human-machine interaction and of the search process is in its infancy [Saracevic, 1987]. This discrepancy is to the advantage of research: Since databases are already in use, investigators do not have to simulate situations in anticipation of the future--they can study online searching behavior as a phenomenon actually occurring in the real world.

The research project reported here investigated online searching behavior manifested by actual searches of bibliographic databases. Despite the growing number of users who search their own requests, the study focused on the process of online searching as performed by human intermediary searchers, that is, by professional online searchers. The study explored the process of search key selection, and attempted to represent this process in an empirically-based model that is specified in formal terms. Such endeavor is an important contribution to research in online retrieval because it enhances our understanding of "what is actually happening at the man-machine interface in online systems" [Fenichel, 1980]. The construction of such a model is valuable to basic research, to the training of online searchers, and to the design and development of information retrieval systems that can be searched by users (other than professional searchers).

1.1 Problem Definition

One of the tasks carried out by information professionals when performing an online search of bibliographic databases is the selection of search keys. To understand the nature of this task it is best to examine its place in the process of online searching.

Online searchers handle requests which reflect the library user's information need. Once the searcher feels he understands the request well enough to answer it, he performs an online search on the relevant databases. Typically, the outcome of an online search is an answer set that includes bibliographic citations.

Reality, however, is much more complex than this description. First, "information need" is an elusive concept: Even if a real and precise need for information exists in an objective sense, it is difficult to determine it accurately. Asking users to define their information needs requires them to describe in exact terms what they do not know--a situation that is most often contradictory in nature [Belkin & Vickery, 1985]. For the purpose of this research project, however, we assumed that what are expressed by users when they want to retrieve information are information needs that are clearly defined.

Further, users' requests have two major aspects. The first is the topic of a request: It presents the subject matter that is of concern to the user. For example, "the analysis of students' behavior during a final examination to determine the difficulty of the examination" is the topic of a hypothetical request.

The second aspect of user requests concerns request characteristics that do not relate directly to the topic but rather to the purpose of the request, or to the use to which the information will be put. For example, a user may need a comprehensive search that retrieves all the relevant citations (high recall search), or she may be interested in just a few highly relevant citations (high precision search). In addition, at one time the user may agree to consider articles about any examinations--not only finals--and at another time she may be happy to receive citations to articles dealing with the analysis of students' behavior during final examinations, whether or not the analysis is used to determine the difficulty of the examination (low specificity search).

With the intricacy of requests exposed, we can turn to the description of the search process. The classic online search includes the following procedure. A searcher interviews a user to clarify the topic and the characteristics of a request. The searcher then develops a plan for searching the request online--a search strategy. This strategy specifies which databases will be searched and which terms (or search keys) will be used in each database. It can also include a more specific plan that determines the flow of the search: Which search keys to enter first, when to review some results, and what to do if the results are not satisfactory. Next, a session at the terminal is guided by the search strategy but searchers may deviate from their original plan if it does not seem useful. Some requests may require a number of terminal sessions--searchers may logoff to reconsider their strategy, possibly with the help of the user. At some point, the searcher decides to terminate the search and to print the answer set that will be given to the user.

Thus, the intellectual components of a typical online search can be classified into three basic categories: (1) definition of query structure; (2) selection of search keys; and (3) feedback review. The second category--the selection of search keys--was the focus of this study.

To select search keys for a request, searchers must first break down a request into its individual components, or concepts. The request about students' behavior, for instance, includes four concepts: (1) analysis; (2) students' behavior; (3) final examinations; and (4) examination's difficulty. Each concept requires a set of search keys for its representation. Thus, searchers look for search keys that will best capture the literature on the topic of each individual concept or of the request as a whole, and at the same time retrieve an answer set that satisfies other request characteristics, such as recall, precision, or specificity.

There are two distinct types of search keys: free-text terms and descriptors. Searchers may enter any desired term or phrase and retrieve citations that include the term or the phrase in their text, that is, perform _free-text searching_. Or, searchers may decide to use search keys from a thesaurus--a list of descriptors, or subject headings that are used for indexing and retrieval--that is, perform _descriptor searching_. In many databases, both options are available. Searchers can also use them in combination: A concept may be searched using free-text search keys as well as descriptors.

One of the decisions that searchers make during an online search, then, is what type of search keys to use; they make this decision when they plan their strategy and during terminal sessions when they revise their strategy. The research project reported here studied the _decisions_ that searchers made when they selected search keys, and it aimed to uncover their reasons for the selection of each _type_ of search keys.

1.1.1 _Controlled Vocabulary and Free-Text Keys_

The issue of search key selection has been the focus of many research projects and publications. As Svenonius points out, the debate over whether controlled vocabulary is necessary for effective retrieval began in the last century, long before the introduction of computers [Svenonius, 1986]. Although this debate originated from problems encountered when using controlled vocabularies with printed catalogs, the notion that controlled vocabularies are an unnecessary burden on information scientists and specialists has been the driving force behind much research in recent years.

On the theoretical front, the construction and use of controlled vocabularies involves a large number of variables, and some fundamental issues have not yet been resolved. For example, there is no agreed-upon measurement for the degree of control exercised in a given index language, nor any well-grounded theories about the factors that constitute useful indexing practice. From a practical point of view, controlled vocabularies are expensive to construct and indexing is labor-intensive as compared to free-text searching because the text is already available and requires only the automated generation of indexes.

Despite the expense and difficulties in the construction of controlled vocabularies, they *are* built and used because they improve retrieval. It is not surprising, therefore, that studies to examine their necessity centered around retrieval performance. Starting with the Cranfield studies [Cleverdon, 1962], investigators have carried out tests to determine which types of search keys provide the best retrieval: free-text terms or descriptors (e.g., [Parker, 1971], [Keen, 1973]). Results are contradictory; the issue is still unresolved and is heavily debated in the literature (e.g., [Cleverdon, 1984], [Lancaster, 1980], and [Dubois, 1987]).

While some may believe that persistent experimentation will eventually resolve the issue of which type of search keys is best for retrieval, there is an increasing evidence that free-text and descriptor searching actually *complement* one another, and no single type outperforms the other. This relationship has been derived by Fugmann from his theory of indexing [Fugmann, 1982], tested in several experiments [Katzer et al., 1982], and substantiated by a series of independent case studies (e.g., [Carrow & Nugent, 1981], [Henzler, 1978], and [Markey et al., 1980]).

The study reported contributes to the resolution of this controversy. The study's purpose was to develop a model that represents the rules for the selection of search keys. The uncovering of such rules would show that each type of search keys is selected for a reason; it would thus prove that free-text and descriptor searching indeed complement one another, but more importantly, the model would show *how* they complement one another.

It is most fruitful to study actual search processes because the selection of search keys is an important component of searching behavior (e.g., [Baker & Eason, 1981], [Oldroyd, 1984]). During online searching, human intermediaries base a large part of their decisions on the tradeoffs between free-text and descriptor searching, and it is one of the most important vehicles for improving search results [Fidel, 1984b]. In addition, searchers' deliberations can be observed and recorded to uncover the hidden and somewhat intuitive rules they use for the selection of search keys.

1.1.2. *Intermediary Expert Systems*

It is believed that an increasing number of users prefer to interact directly with online bibliographic retrieval systems. Although no statistics exist as yet to support this assumption, a large amount of effort is being invested by software producers and search system vendors in developing systems, such as SciMate or Colleague, that facilitate online bibliographic retrieval from users' offices or homes. It is also believed that users will very likely search their own requests online when search processes are simplified or made friendlier. The prevailing approach to providing such user-system communication is to develop intermediary systems which are designed to mediate between users and complex information retrieval systems.

With intermediary expert systems, users should be able to present their requests to a system which would then make expert decisions about the search process and, in particular, about the selection of search keys. Such systems should interrogate users to elicit request characteristics, but a system will use its own expertise to make decisions about matters that are beyond the knowledge of users. For example, an intermediary expert system should ask the user whether high recall or high precision is required, and will use this information to decide whether to use free-text terms or descriptors or a combination of the two.

Various intermediary systems are already available for public access, such as CITE [Doszcoks, 1983], while others are prototype systems being tested in experimental settings. Examples of the latter are CANSEARCH [Pollitt, 1987], PLEXUS [Vickery & Brooks, 1987], EP-X [Krawczak et al., 1987], or CoalSORT [Monarch & Carbonell, 1987], each covering a limited subject domain and searching a single database. Through such systems, users are freed from encounters with the numerous peculiarities of databases and search systems--such as ORBIT, DIALOG, or BRS--and yet can benefit from a large range of capabilities. In particular, an intermediary system allows users to enter a request in a loosely structured format, preferably in natural language, using a sentence-like expression. The system then processes the request terms, displays information to users, and asks for feedback. The information displayed may be in the form of a list of subject areas, databases, search keys, or actual citations from which users are asked to make a selection, possibly in ranked order. Interaction of this nature usually proceeds until the user terminates the session.

Some intermediary systems are actually helper systems: They provide menu-driven interaction that frees users from learning the command language while still requiring them to make most of the decisions during a search process; or, they drastically simplify searching by reducing the number of options to a minimum. CITE, for example, leaves the selection of search keys to the user: It displays a list of search keys that can be used for a request concept--both descriptors and free-text terms--and asks the user to select the terms. In contrast, CONIT--which provides an interface with a number of databases covering a variety of subjects--simplifies the selection of search keys because it searches each search key as a free-text term and, under certain circumstances that depend on the search system rather than the request, also searches each as a descriptor [Marcus, 1983].

Intermediary expert systems, on the other hand, attempt a more powerful form of user assistance: They replicate the performance of an expert in online bibliographic retrieval by incorporating the knowledge of an expert with rules for making inferences on the basis of this knowledge. Well-advanced expert systems are expected to select search keys. In a database that offers both controlled vocabulary and free-text searching, such systems must examine each term of a request and consider its representation as a descriptor key, as a free-text key, or as both.

Intermediary expert systems have attracted much attention and controversy [Smith, 1987]. Although a variety of definitions for expert systems currently exist, most researchers agree about the nature of intermediary expert systems. Studies examining users searching their own requests with no intermediary assistance show repeatedly that users needed intermediary expertise mostly for formulating search strategies, while they seem to master the command language with no difficulties (e.g., [Sewell & Teitelbaum, 1986], [Kirby & Miller, 1986]). Therefore, every intermediary expert system that is being developed today must include a component that supports decisions about search strategies and, in particular, about the selection of search keys.

Daniels [1986], Brooks [1987], and Croft [1987], among others, delineate the requirements for such intermediary systems. One such requirement is that an intermediary expert system should be able to take into account request (and user) characteristics that are beyond the topical description of the search.

This requirement is discussed here because this study is pertinent to its implementation. Although various techniques are used to develop user models [Daniels, 1986], it is not clear what user characteristics are important for the success of an information-retrieval encounter. For example, can the age, profession, or geographic location of a user help an intermediary expert system decide on a search strategy? Paice emphasizes the significance of this point when he observes that unlike other expert systems, user interaction plays a central role in intermediary systems, and the main concern is, therefore, what questions to ask and when to ask them [Paice, 1986]. A model of decision rules used by human intermediaries for the selection of search keys could uncover these questions and suggest a sequence for their display.

Such a model might show, for instance, that while online searchers do not take the age of a user into consideration, they do use their knowledge about the user to determine whether the user's professional activity is focused on practical applications or on research. It might also point out that for some terms the selection of search keys is limited to one option, regardless of request characteristics, while for other terms those characteristics play an essential role in decisions about search key selection.

The importance of human expertise to the design of expert systems is still a controversial issue. It seems, however, that the notion that intermediary systems should be based on knowledge acquired from human experts is gaining increased recognition. Croft, for example, maintains that the formalization of the knowledge used by human intermediaries is one of the open problems of research in expert systems for information retrieval [Croft, 1987], and Daniels mentions it as the most promising method for the construction of user models [Daniels, 1986]. In addition, a few prototypes--such as PLEXUS [Vickery et al., 1987], and EX-P [Smith, et al. 1987]--are already based on such a knowledge.

The research reported here analyzed searching behavior of human intermediaries and then presented this behavior in a formal model. It thus represents the first step in incorporating experience gained by human intermediaries into knowledge bases of intermediary expert systems.

1.2. The Objectives of the Study

To begin a systematic investigation of searching behavior, I first completed a study of online searching behavior using the case-study method [Fidel, 1984a]. Eight experienced human intermediaries were observed doing their regular, job-related searches, and their spoken thought processes were recorded. Analysis of data collected in this preliminary study uncovered three major patterns in searching behavior.

The first delineated searching styles: It described the operationalist and the conceptualist searchers, their approach to strategy formulation, to the selection of search keys, and factors they considered important to decision making [Fidel, 1984b]. Briefly, operationalist searchers aimed at optimal strategies to achieve precise retrieval; they used a large range of system capabilities in their interaction. They preserved the specific meaning of a request, and the aim of their iterations was an answer set representing the request precisely. In contrast, conceptualist searchers analyzed a request by seeking to fit it into a faceted structure. They first entered the facet that represented the most important aspect of the request. Their search was then centered on retrieving subsets from this primary set by introducing additional facets. Unlike the operationalists, they were primarily concerned with recall. During the interaction they preserved the faceted structure, but, if needed, they did not hesitate to change the specific meaning of the request.

The second pattern that emerged from this study represented moves that are made by searchers. It showed that each move belongs either to the set of moves that are typical of an operationalist searcher or those typical of a conceptualist. Operationalist and conceptualist moves were then clearly divided into: (1) moves to reduce the size of the set; (2) moves to increase the size of a set; and (3) moves to improve both precision and recall [Fidel, 1985].

The Selection Routine was the third pattern which emerged. The Routine is a presentation of rules for the selection of search keys in the form of a decision tree [Fidel, 1986].

1.2.1 Refinement and Validation of the Selection Routine

While the Selection Routine clearly indicated that formal rules could be extracted from human experts, it was incomplete at that stage of study. First, there were a number of conditions that led to more than one option. For example, if a term was a common term (that is, not appropriate for free-text searching) and it was not mapped to a descriptor, expert systems were left to decide whether to use free-text terms to probe indexing, or whether to change database. Clearly, there might be additional conditions that would determine which of these options to select, but these conditions were not revealed by this previous study. Secondly, the eight searchers that were observed for that study were experts in the life sciences literature, but to build a Selection Routine that is applicable to bibliographic retrieval in *every* subject area, the searching behavior of human intermediaries in a *variety* of subject areas had to be investigated.

Thus, the first objective of the research project reported here was to refine and validate the Selection Routine.

1.2.2 *The Effect of Searching Behavior on Search-Key Selection*

The description of operationalist and conceptualist styles of searching [Fidel, 1984b], suggested that, among other factors, the selection of search keys might be determined by the searching style of each individual searcher.

Thus, the second objective of this research project was to test the hypothesis that searching style affects the selection of search keys, and to uncover the nature of this effect.

1.2.3 *The Applicability of the Case Study Method*

The case study method with controlled comparison [Diesing, 1971], which was used to generate the Selection Routine, offers solutions to some of the problems in knowledge acquisition. First, it has two useful attributes: (1) it adds knowledge incrementally; and (2) it is equipped to resolve contradictions arising from the use of more than one case. Second, this study method is a systematic and well-structured approach to extracting knowledge from human experts.

If the use of the case study method proved useful in one research project, it can be transferred to other knowledge domains. It is a method that can be used to extract knowledge from a number of experts and to incorporate new knowledge into existing systems.

Thus, the third objective of this research project was to test the applicability of the case study method to the extraction of knowledge from multiple experts.

2. THE METHOD

The case study method with controlled comparison [Diesing, 1971] was used to investigate the selection of search keys. Briefly, in this method a case is analyzed to construct a model of the investigated phenomenon based on one case. An additional case, which is similar in a definite sense to the first case, is then analyzed and is fitted into the model created by the first one. Discrepancies are resolved either by increasing the level of generality in which the elements of the model are expressed, or by adding elements to the model. The modified model of the investigated phenomenon is now based on two cases. Additional cases are analyzed, one after the other and representing a gradual increasing diversity, to further refine the model and to expand its applicability. Models constructed by the case study method with controlled comparison are never complete (in an absolute sense): The more cases are analyzed, the more general the model is. They are dynamic, however, in that they can be modified and expanded to fit new developments and discoveries in the investigated phenomenon. A detailed description of the use of this method in the investigation of online searching behavior is available elsewhere [Fidel, 1984a].

The data for this project were collected through observation and interviews.

2.1 Data Collection

To collect data, members of the research team observed professional online searchers when they were doing their regular, job-related searches. The searchers were asked to think out loud as they worked. Their verbalization was recorded and transcribed, and together with the written material (e.g., the search protocol, the request form) served as a basis for the analysis, which was primarily protocol analysis.

Protocol analysis was used in the project as a first attempt to clearly define each move in a search, and to identify and analyze each instance in which a search key was selected. Once such an instance was established, verbalizations of thought processes, previous and proceeding moves in the search, and recorded search strategy were used to explore the conditions that had led to the particular selection.

Members of the research team interviewed each searcher immediately after the sequence of observations for that searcher had been completed. Before an interview was conducted, all the transcribed protocols of searches performed by a searcher were analyzed to identify issues that were inaccessible to observation or those that needed clarification. In the interviews, searchers were asked to explain their moves and reasons for selecting individual search keys. The interviews were then transcribed. Answers of searchers were checked for validity by comparing them with other types of evidence.

2.1.1 The Procedure

The research team included the principal investigator aided by two research assistants at a time, who were selected from among second-year students at the School of Library and Information Science at the University of Washington. The research assistants (a total of four students) performed the observations and conducted the interviews. Each assistant--or observer--investigated one searcher at a time. Observers were trained by actually performing all the field investigations and analyses for a selected searcher. Although these four searchers were initially selected only for the purpose of training, data collected from their searches were eventually incorporated into the project because of the high quality of the outcome of these training activities.

The procedure of data collection included the following steps: (a) contacting the searcher; (b) recording sessions; (c) transcribing recorded material; (d) analyzing search protocol and verbalizations; (e) recording data on forms; (f) participating in weekly meetings to report findings; and (g) interviewing the searcher for clarifications.

(a) Contacting the Searcher.

The principal investigator made the initial contact with most searchers. She explained the purpose of the study and briefly described the commitment that was required from the searcher. The investigator explained that the purpose of the study was to understand how online searchers perform their searches; it was an attempt to make the "art of online searching" more explicit and available to scientists and system designers as well as to new searchers.

Searchers were promised that almost no demands would be put on their time and were asked to briefly explain what they were doing so the observer could understand the search. The investigator also clarified that because the observer might forget the searcher's explanation, the observer would have to record these explanations on tape, but that the recordings would be used only by the research team. After repeated requests, the investigator also promised searchers that the results of the study would be made available to them as soon as possible.

Following the initial contact, an observer met with an assigned searcher for a brief introductory session. In this session, the observer further explained the searcher's role, established a convenient method for communication and for setting future appointments. In addition, the observer inquired about the databases the searcher used most frequently so that the observer could get familiar with the databases and the search system used.

The observers were asked to be honest when communicating with searchers and to hide no information from them. Obviously, to protect the privacy of the searchers, no information about their searching behavior was given to other searchers. However, information about the study, its progress or difficulties, was available to searchers if they were interested.

The observers were also instructed to be flexible and non-demanding when they set up appointments. They told each searcher that they would operate according to the searcher's schedule, and established a method by which either the observer would call the searcher periodically, or the searcher would call the observer when planning to perform a search. In addition, the observers checked whether scheduling would permit the observation of the reference interview. In most instances, however, such an observation could not be arranged, with one important exception: A number of the observed searches were conducted with the user present at the terminal.

From the beginning of the observation period it was explained to searchers that any idea or notion that crossed their minds while searching would be of importance to the study. While data analysis focused on reasons for the selection of search keys, it was important that searchers did not pay attention to this issue in order to avoid bias or influence on searching behavior. The focus of the study was later revealed to the searchers during the interview so they could explain their reasons for the selection of particular search keys.

(b) Recording Sessions

During the observation, the observers tried to be non-disruptive, non-threatening, and non-judgmental. To help them in this pursuit, they were constantly reminded that they were observing experienced and professional searchers who were more knowledgeable in the practice of online searching than any member of the research team. Experience in the study shows that having students perform the observation and interviews was advantageous in this respect because it was natural for searchers to expect non-threatening and curious behavior from students who were not yet professionals.

Working with student observers on this project provided another important advantage. As future professionals they were eager to learn about online searching behavior. Their natural curiosity made them ideal listeners: They were genuinely interested in the phenomenon they were investigating. This curiosity also motivated them to perform the analyses of search protocols and transcriptions.

Questioning during the observation was kept to a minimum so that searching behavior would not be affected, and to eliminate any additional burden on the searcher. However, when an observer missed one of the steps, or when she did not understand what the searcher was doing, she would ask the searcher to describe the steps that were taken. All questions about reasons for taking a particular step were saved for the interview which took place at the end of the observation period.

Because most often the observers could not observe the reference interview, they began each session with a request that the searcher explain the topic of the request and its nature. The observers also made copies of the search protocol, the request form, or of any piece of paper that was relevant, such as the paper on which the searcher had recorded search strategies or moves.

(c) Transcribing Recorded Material

All the verbalizations during a session, whether directly related to the search or not, were transcribed. This approach was adapted because previous experience had indicated that some pieces of information that seemed irrelevant at one point turned out to be of great importance at a later time. The verbalizations were typed consecutively in the form of a dialog, with references to search statements.

(d) Analyzing Search Protocol and Verbalization

To provide a view of the whole search, the "history" of each search was sketched, recording each formulation entered and the number of the resulting citations (postings) for each search statement.

The first analysis uncovered the moves made. Here the observers examined each search statement and compared it with the previous ones to detect any change in the search strategy. Once a change in strategy was detected, the observers examined the list of moves [Fidel, 1985] to identify the nature of the move. For example, if a searcher entered in one statement the formulation "evaluation AND methodology," and the next statement was "(evaluation OR assessment OR determination) AND methodology," the observer noted that the move ADD 1 was made (i.e., add synonyms and variant spelling).

Moves that were detected but could not be identified using the list of moves were discussed at the next meeting of the research team.

The second analysis examined the selection of search keys. For each instance in which a searcher selected a search key the observer examined the Selection Routine in order to identify the set of conditions that resulted in that particular selection [Fidel, 1986]. For example, if a user asked for material about mini-pigs and the searcher entered the descriptor Swine, the observer denoted the option "J" (when the term is a single-meaning term and it is mapped to a broader descriptor--enter descriptor).

In the analysis of the selection of search keys the observer also recorded the reason for the decision to select a particular search key. In the above example, the observer might record that the reason for the selection of a broader descriptor was the user's requirement for high recall.

Instances in which the selection of a search key did not correspond to any option in the Selection Routine were discussed in the team's meeting. If the reasons for a selection of a search key were unclear to the observer, she would make a note to ask the searcher later--during the interview--for additional clarification. The validity of such retrospective answers is discussed in section 2.1.2 (Adjustments).

(e) Recording Data on Forms

Data were recorded on three forms: the Moves Form, the Selection Form, and the Reason Form. These forms are presented in Appendix A.

(1) The Moves Form consisted of the list of moves arranged by their purpose (to increase the size of a set, to decrease the size of a set, or to improve both precision and recall), and by the type of the move (operational or conceptual). One Form was used to record data collected from one search, noting the name of the searcher, the name of the search, the number of times each move was made, and the search statement in which each move was made.

At the end of the observation period, another Moves Form was used to summarize the moves made by a searcher during all the observed searches. The Form itemized the name of the searcher, the number of times each move was made, and the name of the search in which each move was made. In addition, the number of operational moves, the number of conceptual moves, the total number of moves, as well as the percentage of operational and conceptual moves was recorded at the bottom of the summary Form.

(2) The Selection Form consisted of a list of all the conditions and options as presented in the Selection Routine. For example, the letter "A" represented the option: When the term is a common term (not adequate for free-text searching) and it is mapped to a descriptor--use descriptors. In the same manner, the letter "H" represented the option: When the term is a single-meaning term and it is mapped to a broader descriptor--use free-text terms.

One Selection Form was used for each searcher. On it, observers recorded the name of the searcher, the number of times each option of a search-key selection had occurred, and the name of the searches in which it occurred.

(3) The Reason Form was designed to record the reasons for selecting a particular search key. Each option from the Selection Routine had its own form. Thus, observers recorded reasons for the "A" option on one Form and those for the "H" option on another Form. On each Form, the observers recorded the identity of the searchers, the search names, the search keys themselves, and the reasons for the selections.

In addition, to facilitate integration of data recorded on the various forms, and to provide more detailed information about the selection of search keys, each selection option had its own card. On the card, the observers recorded the search names, the searchers' names, the search-statement numbers, and the search keys selected.

(f) Participating in Weekly Meetings

The research team met about once a week. In these meetings each observer described the searches she had observed during the week and presented both the moves she had detected and her analysis of the selection of search keys. The team would then discuss the search, examine the observer's analysis, and identify issues that required further clarifications from the searcher. In that respect, the entire team was acting as a panel of judges.

In addition, these meetings served as a vehicle for observers to share their new discoveries. An observers would always bring a newly detected move, or an unanticipated option for the selection of search keys before the research team. Such new discoveries would not be added to the model before the entire team had discussed them. Obviously, through these meetings each member of the team was immediately informed about modifications in the list of moves or in the Selection Routine.

Another important function of the team meetings was to devise policies about interpreting ambiguous observations. Such policies addressed both general interpretations of searching behavior and interpretations of specific moves or search-key selection.

The research team decided, for instance, that only subject-related search keys would be considered as search keys. Another example of a general policy stated that if a searcher saved a search and then executed it automatically in another database no selection of search keys would be recorded, but if a searcher re-entered a search formulation when searching a new database the search keys would be considered as newly selected.

An interpretation for the use of truncation is an example of a policy decision that related to a specific move. The team decided to record each truncation as representing the move ADD 1 (i.e., add synonyms and variant spellings). Similarly, the team decided to record the move LIMIT 1 (i.e., limit to documents written in a particular language) only when it was applied to reduce the size of the set, and to ignore it when it was applied because the searcher believed that the user wanted documents in English only.

Further, the discussions held during the team meetings prepared the observer for the final interview with the searcher.

(g) Interviewing the Searcher

Before interviewing a searcher, the observer and the investigator discussed the questions to be asked, their language, and the type of information to be elicited from the searcher. The observer then met with the searcher and later transcribed the interview. The clarifications provided were then discussed at the next team meeting.

All the data collected for this project were kept in a locked office. To protect the privacy of the searchers who participated in the study, only members of the research team had access to these data.

At the end of her term with the research team, one of the observers wrote a description of the procedure followed, noting technical problems and suggesting modifications. This write-up was used in the initiation of new observers and it is given in Appendix B.

2.1.2 Adjustments

Typical of a field research method, the case study method is very time-consuming. Data collection and analysis were labor-intensive. In this particular study, because searches were recorded as they naturally occurred, the research team was unable to impose schedules or deadlines, thus waiting for a case was time consuming. We had to abide by the searchers' schedule which at times did not agree with our plans. This situation created an occasional need for adjustments as described below.

(a) The Selection of Cases

As explained above, the case study method with controlled comparison requires that each case be similar to the previous ones in a definite sense. This implies that cases (i.e., searchers) are selected in a certain sequence--a sequence that is determined by the searchers' characteristics, such as the type of library in which they work, or their subject specialty.

This principle of selection was followed only during the preliminary preparation for the project, and it proved to be crucial for the success of that stage [Fidel, 1984b]. Time constraints, however, prevented us from following any ordering principle in the selection of searchers for the project itself. Because the number of searchers to be observed was relatively large (40), considering the time allocated for the study (2 years), we observed searchers in the order that they made themselves available.

As the study progressed, however, we became more careful in the selection of searchers; we aimed at equal representation of searchers of each type. For instance, after the first year it became clear that we would have difficulties in recruiting "enough" science searchers from academic libraries. We then expanded our territory to include searchers from out of state and successfully engaged a number of science searchers.

Although a deviation from the method, this change in the procedure did not create any obstacles. In retrospect, it seems that the study was not affected by this change because both the list of moves and the Selection Routine were already well-structured. Our experience leads, therefore, to the conclusion that while the order in which cases are selected is of paramount importance for the beginning stages of model construction, its role is not always critical for the expansion of a mature model.

(b) Retrospective Observations

The observation period for each searcher varied greatly, ranging from a week to almost three months. After the first year we encountered a number of searchers whose work patterns made direct observation impractical. One group of such searchers were special librarians working in one-person libraries, who could not schedule their searches ahead of time: Most often they needed to perform a search immediately after receiving a request. The other group were out-of-state searchers who could not complete enough searches during the period of time the observer visited their location.

To avoid the under-representation of special librarians and out-of-state searchers in the study's sample, we decided to interview these searchers after they had performed their searches. Special librarians called the observer immediately after they had performed a search and set an appointment at their earliest convenience. Out-of-state searchers who did not believe they could furnish enough searches were notified ahead of time about the dates in which the observer would visit their location and were asked to keep their last searches. The searchers prepared copies of the search protocols, and of any other relevant documents. They then described each search to the observer, explaining their strategies, search-key selections, and moves. These descriptions were recorded and transcribed.

This adjustment in the study procedure had advantages and limitations, as explained by Ericsson & Simon [1984]. The most significant advantage was that searchers were in a way forced to explain their searches. Eliminating the time pressure present during an online search, searchers could describe in great detail each step they had taken. In addition, observers could ask searchers to clarify their reasons for search-key selections and moves as the searches were described. Describing searches after the fact, searchers were also willing to explain all the searches they had performed in a recent period. As a result, they often provided information about more searches than could be observed. The most prominent limitation of this approach stems from the fact that searchers might not recall the dynamics of a search. We found that searchers had no difficulties reconstructing their previous searches. One aspect of the search process was missed, though: For searches done with the user at the terminal, searchers could not always recall who had suggested new terms or new strategies--the user or the searcher.

One of our immediate concerns was searchers' inability to retrace their thought processes *accurately*. After all, one could not expect them to remember the reasons for every decision they had made during a search. After much discussion we concluded, however, that this concern is not completely relevant to the research project. The purpose of this project is to uncover the intuitive rules that online searchers employ when they select search keys, and whether a particular rule was applied in a specific search is of little concern. We would be concerned about this modification if we were aiming at finding which rules are *actually* used by each individual searcher for each individual search.

We assumed, therefore, that if a searcher stated a rule for a selection of a search key, this searcher employed this rule; whether or not this rule was indeed employed during the *particular search described* was not significant. Consider for example a search reconstruction in which a searcher explained that he had selected a broader descriptor because the user was interested in high recall. The data that were pertinent to the study is that high-recall requirement was a reason *used by that searcher* for the selection of a broader descriptor. Whether the searcher indeed used this rule for the particular search he described was not critical to our project.

This adjustment was employed for a total of nine searchers.

(c) Initial Contact with Searchers

At the beginning of the project the investigator was the only person to create the initial contact with the searchers. As the study progressed, and the observers became familiar with a number of searchers, it became common for a searcher to recommend another searcher. Since the observers knew the source searcher, and because they felt knowledgeable enough to introduce themselves to new searchers, they would contact the new searcher. Thus, in a number of cases the initial contact was made by the observers.

(d) Co-observation

Initially, each observer observed only one searcher at a time. After the first year this practice was changed for two reasons. First, we decided that we can no longer afford to wait for "slow" searchers. Therefore, if a searcher did not schedule appointments frequently enough (at least one appointment in three weeks), the observer took on another assignment. Second, the time the observers could spend on trips was limited because of their school schedule. As a result, a number of out-of-state searchers were observed during one visit.

2.2 The Selection of Searchers

To accomplish the objectives of the research project, 39 experienced online searchers were selected for observation from among information analysts--30 of whom work at the Puget Sound area and nine from California. Eight searchers agreed to participate in the project but withdrew after one or two observation sessions, primarily because of scheduling difficulties. Each searcher was asked to allow observation for five searches. While most searchers were actually observed for five searches, the number of searches per searcher, in a few cases, varied from four to eight.

Subjects who qualified for participation were searchers who had been searching for at least two years, and ordinarily searched databases that provide both free-text and descriptor searching. At the beginning of the project, only searchers who performed an average five searches a week were selected. This requirement was introduced to facilitate the selection of subjects who were indeed active searchers. After a short period, however, it became clear that this requirement is too restrictive: Some searchers could not specify the average number of searches per week because of great fluctuations in their work load, and highly experienced and active searchers who worked part time did not always complete five searches a week. This requirement was, therefore, dropped.

To improve the generality of the model, searchers were selected from a wide spectrum of subject specialties. The sample included 16 searchers from the humanities and social sciences, 21 from the science and technology area, and three medical librarians.

The selection of subjects for the study was guided by the following considerations.

The number of searchers to be observed was determined by the number of searches needed for the analysis and by the number of observations each searcher could tolerate. There was no method that could be used to provide an estimate of the number of searches that were needed for the project. However, based on the preliminary study, in which about 80 searches were analyzed [Fidel, 1986], it seemed that 200 additional searches would need to be analyzed in order to resolve most of the ambiguities in the Selection Routine. This number of searches was also large enough to accommodate the variety of subject areas in literature searches, and it seemed to provide a sufficient sample of search-key selections to test the suggested hypothesis. Previous experience of the investigator with field observation indicated that online searchers can comfortably tolerate observation for five searches. The number of actual searches analyzed for this project is 201.

The qualifications for the subjects were selected because of the following considerations: (1) subjects had to be experienced online searchers. The requirement that they had two years of experience was based on the assumption that the amount of experience gained during this period of active searching had crystallized their searching behavior. (2) Searchers who searched databases that facilitate searching with only one type of search key are rarely in a situation in which they must decide about a desired type of search key and therefore are not likely to develop rules for search-key selection. The experience of such searchers could not contribute to the Selection Routine. In addition, searchers with various subject specialties were chosen to develop a Selection Routine that is applicable to any subject area.

2.3 Data Analysis

Typical of a field study and of using a qualitative method, data analysis was carried out from the very beginning of the project. Unlike surveys or experiments where data are first collected and only then analyzed, the case study method requires ongoing data analysis. Indeed, some aspects of data analysis were already described when the procedures followed were explained (section 2.1.1).

Miles and Huberman [1984, p.21-23] point out that data analysis in qualitative research has three components:

(1) data reduction which selects, focuses, simplifies, and abstract the "raw" data--organizing them in a fashion that would facilitate conclusion drawing and verification;

(2) data display which assembles organized information in an immediately accessible, compact form so that the analyst can see what action is needed next; and

(3) conclusion drawing/verification--noting regularities, patterns, explanations, possible configurations, causal flows, and propositions.

They further explain that "the three types of analysis activity and the activity of data collection itself form an interactive, cyclical process. The researcher steadily moves among these four 'nodes' during data collection, then shuttles among reduction, display, and conclusion drawing/verification for the remainder of the study." [p.22]

Data reduction was performed in this project when individual searches were analyzed to identify moves and options in the selection of search keys. Data display was carried out by presenting the Selection Routine in the form of a decision tree, and by listing the moves in a table according to the purpose of each move and its type, operational or conceptual.

Conclusion drawing/verification was finally completed to satisfy the objectives of the study.

First objective: To refine and validate the Selection Routine.

Search protocols were systematically analyzed, one after the other, to identify incidents where a search key was selected. Each such incident was then fitted into the decision tree, following the method of controlled comparison [Diesing, 1971].

The method of controlled comparison facilitated constant modifications of the Selection Routine. The conditions for search-key selection in an incident were determined by data gathered during observation. These conditions were then matched with the equivalent set in the decision tree. Three results were possible: (1) An incident exactly matched a combination of conditions and a resulting option in the decision tree. If a combination resulted in only one option (such as option A in the original SR), the incident did not modify the decision tree but reinforced the rule. If it resulted in more than one option, the reason for the selection of a particular option for the incident, and for the rejection of others, were explored with the searcher in the interview. (2) An incident exactly matched a combination of conditions but did not match any of the options suggested by the decision tree. The selected option was added to the Selection Routine and the reason for its preference over the other options were recorded. (3) An incident did not match any combination of conditions. The new combination of conditions, and the resulting options, were added to the Selection Routine. Previous instances where the same option had been selected were checked against the new combination.

Reasons for electing a particular option were recorded for each option of a multi-option combination. An example of such a combination is the case where a single-meaning term is mapped to a broader descriptor--a combination that suggests three options: use free-text terms, use free-text terms in combination with descriptors, or use descriptors. The reasons were analyzed to determine their nature: whether they were determined by the database, request characteristics, or by the searcher's general beliefs about searching. This analysis, in turn, further refined the set of conditions for each option. In addition, the number of times each combination of conditions resulted in a particular option was recorded to discover which combinations of conditions occur most frequently and which are the most commonly selected options.

At the end of this analysis, the Selection Routine reflected searchers' selection of search keys during 281 searches, performed by 47 searchers (eight searchers who were observed during the preliminary stage and 39 searchers who participated in the project). The modified Selection Routine is presented in chapter 3 (Online Searching Behavior).

Although the Selection Routine is never completed, the number of new combinations and options decreased rapidly as the project progressed. Therefore, it seems that analysis of additional searches would not modify the Selection Routine substantially.

In total, four new options to existing combinations, and one new combination with three options were added to the Selection Routine. While the new combination was popular among the project's searchers, the four new options were each used very few times.

Second objective: To test the hypothesis that searching style affects the selection of search keys, and to uncover the nature of this effect.

Statistical tests were conducted to examine associations between the percentage of free-text terms selected by a searcher (the Search-Key Ratio) and the percentage of operational moves (the Moves Ratio). These variables were also checked against the number of search keys selected for a search, the frequency with which a thesaurus was not consulted, the databases searched, the frequency of changing databases, the subject area, and the environment. Results of these tests are presented in chapter 4 (Factors Affecting the Selection of Search Keys).

Third objective: To test the applicability of the case study method to the extraction of knowledge from multiple experts.

The applicability of the method was tested by its actual utilization. The use of the method is considered successful because (1) the analysis of each additional case added an increment of knowledge and no gaps in knowledge occurred, and (2) the contradictions that arose from the analysis of multiple cases were successfully resolved. As a result, the Selection Routine was refined and expanded to provide a formal representation of rules for the selection of search keys.

2.4 Advantages and Limitations of the Method

One of the reasons research in online searching has not made much progress during the last two decades is the lack of understanding of the search process itself [Fidel, 1988]. Results of studies--particularly of experiments--cannot be explained because the reasons underlying searchers' decisions are not known to the investigators. Employing the case study method in this study is advantageous, therefore, because it supports in-depth analyses, it provides for flexibility, and it facilitates the gathering of data from a variety of sources.

Additional advantages, as well as limitations, of using the case study method in studying online searching behavior are discussed elsewhere [Fidel, 1984a]. In addition, verbal protocols and their use as a source of data are critically viewed in a monograph dedicated to this subject [Ericsson & Simon, 1984]. Here, we briefly point out three limitations that are mentioned most frequently, and the measures we took to overcome their effects.

First, data can be gathered only as events occur, and an investigator cannot direct the course of events to a desirable conclusion. For example, it may happen that ten consecutive searches present the same conditions and in effect do not add new information. The method of controlled comparison helped us to overcome this problem because in it each case that was selected was slightly "different" from the previous ones. In addition, we selected a large enough number of searches to secure the desired variability.

Second, searchers may not be able to articulate reasons for a particular decision about search key selection. We were successful at times in eliminating this limitation by analyzing data from variety of sources--an analysis which put us in a better position than the searcher to identify factors considered in the selection process. In a few instances, we were unable to discover the reasons for a specific decision.

Third, not all data are accessible to observation. This limitation is most typical in protocol analysis, and to compensate for its drawbacks, we conducted an interview at the end of the observation period with each searcher.

3. ONLINE SEARCHING BEHAVIOR

Modifications of the Selection Routine and of the list of moves resulted in updated versions of both models. The Selection Routine was expanded to include both new combinations and new options as well as refinements of conditions that resulted in more than one option. The modification of the list of moves, on the other hand, included only the addition of a few moves that had not been detected before.

This section describes the revised Selection Routine and the updated list of moves. It also provides descriptive statistics about the frequency with which each option and move were selected. These descriptive statistics alone provide new data about searching behavior.

3.1 The Selection Routine

The number of new options that were added to the original Selection Routine is not large. Although this initial Routine has already been described in detail [Fidel, 1986], the refinements of previously-recorded conditions require a complete description of the revised version here.

The modified Selection Routine is presented in the form of a decision tree in Figure 1. This Figure, however, includes neither the refinements introduced nor the frequency in which each condition was encountered. The refinements are presented in the description of the Selection Routine which follows, and the frequency figures are presented in section 3.2 (Searchers' Selection of Search Keys). Table 1 lists the options in the Selection Routine and the associated conditions.

As Figure 1 shows, the first criterion for decisions about the selection of search keys is whether a term is a common term or a single-meaning term. A single-meaning term is a term which is "good" for free-text searching. It usually occurs in a particular context, it is uniquely defined, and it is specific to the concept it represents. A common term, on the other hand, is a term that is not suitable for free-text searching. Such a term usually occurs in more than one context.

For example, in the request about the analysis of students' behavior during final examinations, the terms "students" and "final" are single-meaning terms. By contrast, "analysis" and "examination" are common terms because they may represent different concepts in different contexts. To be more specific, the term "examination" can occur in a subject-related context ("the best way to take student examinations"), being synonymous with "tests." It can be used to represent the concepts of "perusal" or "study" ("examination of students' responses"), in which case the term "examination" could appear in titles and abstracts of articles that are about other subjects. Further, it can be used very loosely to represent the concept of an inquiry of any kind.

The second criterion for the selection of search keys is whether or not a term is mapped to a descriptor. A searcher maps a term to a descriptor (or to a combination of descriptors) when she has decided that a particular descriptor (or a combination) best represents a request term, whether or not there is an exact match between the term and the descriptor. This criterion generates three conditions: a term is mapped to a descriptor, a term cannot be mapped to a descriptor, and the searcher does not know if the term can be mapped.

These two criteria--whether a term is a single-meaning or a common term and whether or not it is mapped to a descriptor--are central to the Selection Routine because they deal with the relationship between concepts and terms: The concepts that need to be represented and the terms that can express them. Since controlled vocabularies are designed to resolve problems in expressing concepts in query formulations, it is important to examine these relationships when analyzing the selection of search keys. This does not imply, however, that these two criteria are

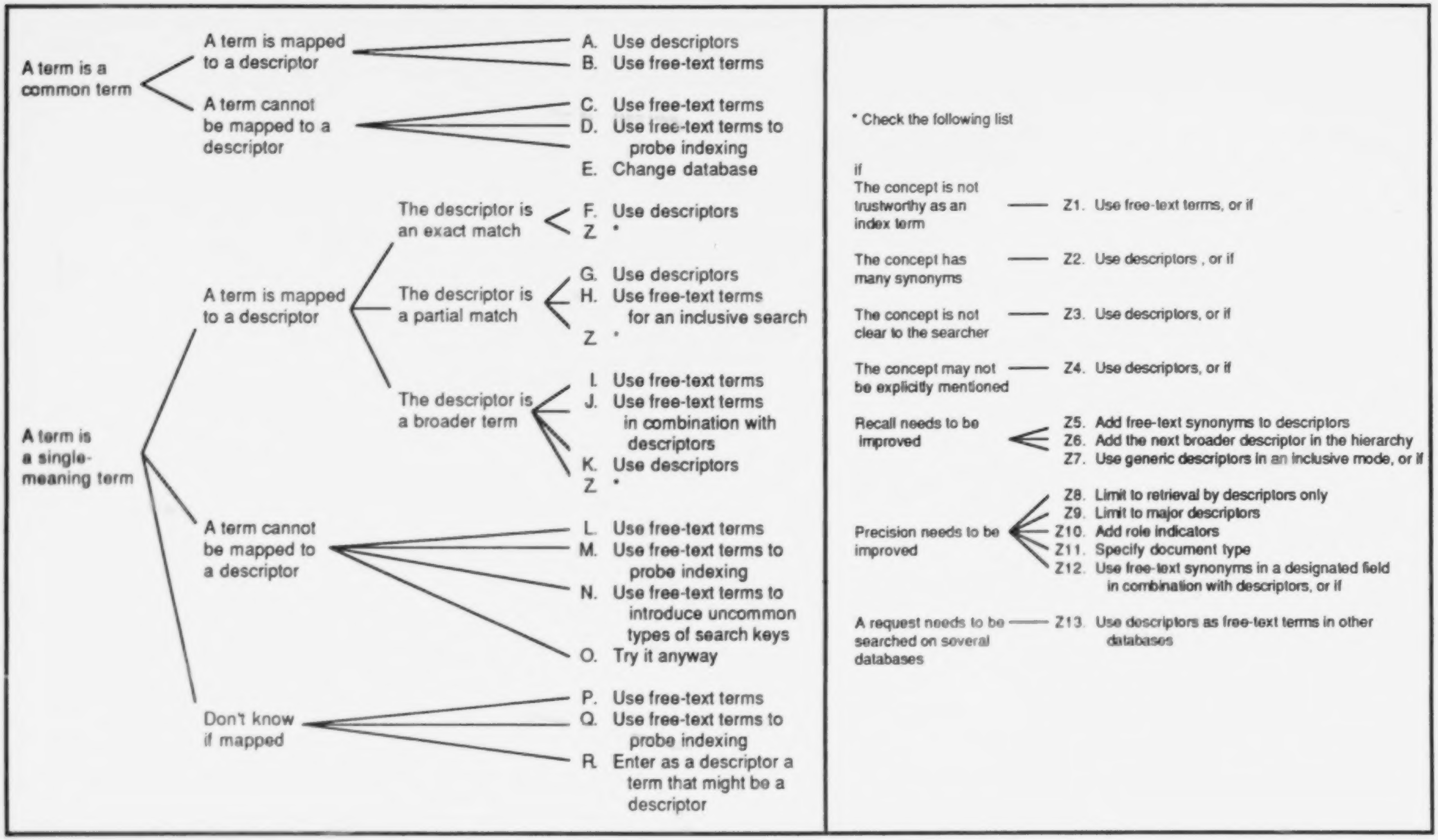
A term is a common term
A term is mapped to a descriptor
A. Use descriptors
B. Use free-text terms
A term cannot be mapped to a descriptor
C. Use free-text terms
D. Use free-text terms to probe indexing
E. Change database
A term is a single-meaning term
A term is mapped to a descriptor
The descriptor is an exact match
F. Use descriptors
Z *
The descriptor is a partial match
G. Use descriptors
H. Use free-text terms for an inclusive search
Z *
The descriptor is a broader term
I. Use free-text terms
J. Use free-text terms in combination with descriptors
K. Use descriptors
Z *
A term cannot be mapped to a descriptor
L. Use free-text terms
M. Use free-text terms to probe indexing
N. Use free-text terms to introduce uncommon types of search keys
O. Try it anyway
Don't know if mapped
P. Use free-text terms
Q. Use free-text terms to probe indexing
R. Enter as a descriptor a term that might be a descriptor
* Check the following list
if
The concept is not trustworthy as an index term
Z1. Use free-text terms, or if
The concept has many synonyms
Z2. Use descriptors , or if
The concept is not clear to the searcher
Z3. Use descriptors, or if
The concept may not be explicitly mentioned
Z4. Use descriptors, or if
Recall needs to be improved
Z5. Add free-text synonyms to descriptors
Z6. Add the next broader descriptor in the hierarchy
Z7. Use generic descriptors in an inclusive mode, or if
Precision needs to be improved
Z8. Limit to retrieval by descriptors only
Z9. Limit to major descriptors
Z10. Add role indicators
Z11. Specify document type
Z12. Use free-text synonyms in a designated field in combination with descriptors, or if
A request needs to be searched on several databases
Z13. Use descriptors as free-text terms in other databases

FIG. 1. The Selection Routine.

Table 1. A list of options and the associated conditions

OPTION	CONDITIONS
	Descriptor Searching
Use descriptors	A term is a common term + it is mapped to a descriptor [A]. A term is a single-meaning term + it is mapped to a descriptor + the descriptor is an exact match [F]. the concept has many synonyms [Z2]. the concept is not clear to the searcher [Z3]. the concept may not be explicitly mentioned [Z4]. the descriptor is a partial match [G]. the descriptor is a broader term [K]. it cannot be mapped to a descriptor [O]. it is not known if mapped [R].
Add the next broader descriptor in the hierarchy	A term is a single-meaning term + it is mapped to a descriptor + recall needs to be improved [Z6].
Use generic descriptors in an inclusive mode	A term is a single-meaning term + it is mapped to a descriptor + recall needs to be improved [Z7].
Limit to retrieval by descriptors	A term is a single-meaning term + it is mapped to a descriptor + precision needs to be improved [Z8].
Limit to major descriptors	A term is a single-meaning term + it is mapped to a descriptor + precision needs to be improved [Z9].
Specify document type	A term is a single-meaning term + it is mapped to a descriptor + precision needs to be improved [Z11].
	Free-Text Searching
Use free-text terms	A term is a common term + it is mapped to a descriptor [B]. it is not mapped to a descriptor [C]. A term is a single-meaning term + it is mapped to a descriptor + the concept is not "trustworthy" as an index term [Z1]. the descriptor is a broader term [I]. it cannot be mapped to a descriptor [L]. it is not known if mapped [P].
Use free-text terms to probe indexing	A term is a common term + it cannot be mapped to a descriptor [D]. A term is a single-meaning term + it cannot be mapped to a descriptor [M]. it is not known if mapped [Q].
Use descriptors as free-text terms in other databases	A term is a single-meaning term + it is mapped to a descriptor + a request needs to be searched on several databases [Z13].
Use free-text terms for an inclusive search	A term is a single-meaning term + it is mapped to a descriptor + the descriptor is a partial match [H].
Use free-text terms to introduce uncommon types of search keys	A term is a single-meaning term + it cannot be mapped to a descriptor [N].
	Other combinations
Use free-text terms in combination with descriptors	A term is a single-meaning term + it is mapped to a descriptor + the descriptor is a broader descriptor [J].
Add free-text synonyms to descriptors	A term is a single-meaning term + it is mapped to a descriptor + recall needs to be improved [Z5].
Add role indicators	A term is a single-meaning term + it is mapped to a descriptor + precision needs to be improved [Z10].
Change database	A term is a common term + it cannot be mapped to a descriptor [E].
Use free-text synonym in a designated field in combination with descriptors	A term is a single-meaning term + it is mapped to a descriptor + precision needs to be improved [Z12].

always used by searchers first and before they examine other factors, such as the constraints of the request or of the database. The priority given to criteria used in the selection of search keys may be determined by the nature of each request, or by the searcher's individual preferences. The question of priority was not examined in this study.

The Selection Routine, as presented in Figure 1, represents only terminological considerations. Searchers who participated in the study, however, considered other factors, particularly for combinations which provided more than one option. These factors fell into three categories: request-related, database-related, and searcher-related considerations. The last category reflected general rules or assumptions that were habitually made by an individual searcher.

We turn now to the description of the Selection Routine.

3.1.1 A Term is a Common Term

A common term may or may not be mapped to a descriptor.

(a) When a common term is mapped to a descriptor. When a common term is mapped to a descriptor, searchers do not have much choice in the selection of search keys: They almost always enter the descriptor as a search key [A] (i.e., option [A] in Figure 1.) because, by definition, it is not desirable to use a common term as a free-text search key.

There is one exception to this rule: Depending on the request, searchers may decide to enter the term as a free-text key [B]. The only instances in which searchers selected this option were when the term was used as a limiting factor, and they perceived that a descriptor might be too restrictive. For example, in the request about the analysis of students' behavior during final examination, a searcher combined the terms "students' behavior" with "final examinations," using the AND operator. Adding the requirement that all citations be also indexed under the descriptor "analysis" might be too limiting, and the searcher decided to retrieve citations that included the term "analysis" in their titles or abstracts--a somewhat less restrictive requirement.

(b) A common term cannot be mapped to a descriptor. A common term that cannot be mapped to a descriptor almost always results in unsatisfactory retrieval. Searchers, however, have almost no choice but to enter a free-text key [C]. Although searchers can enter such a free-text term just to check the indexing of relevant articles, two reasons were cited for a direct use of a common free-text term. The first related to the request and the second related to the database searched. First, if a request includes a relatively large number of concepts--that is, the Boolean operator AND occurs more than twice or three times in the query formulation--precision will not suffer if a common term is entered as a free-text term [CR1] (see Table 2). Second, if a request will be searched on a number of databases, it might be too costly to probe the indexing in each database [CD1].

If a request requires searching only one or two databases, however, searchers can enter the free-text term to probe indexing [D]. One method of probing the indexing is to enter the free-text key in combination with other search keys, in order to retrieve citations, to select some relevant ones, and to review their indexing in an attempt to find descriptors that might possibly be relevant. For example, if the term "examination" cannot be mapped to a descriptor, one can devise a formulation (using the AND operator) that combines the descriptors "students," "analysis," and the free-text terms "final" and "examination." Reviewing a sample of retrieved citations, one may find that all the relevant citations include the descriptor "instructional tests," thus suggesting that this descriptor is an appropriate choice for the representation of the concept "examination."

Such probing does not always further the search and searchers may then decide to select a different database: one which does allow the common term to be mapped to a descriptor [E].

3.1.2 A Single-Meaning Term That is Mapped to a Descriptor

When a single-meaning term is mapped to a descriptor, it can be mapped through an exact match, through a partial match, or to a broader descriptor.

(a) When the descriptor is an exact match. The most direct use of a descriptor to represent a single-meaning term is when a term is exactly matched with a descriptor and no other apparent constraints exist [F].

(b) When the descriptor is a partial match. Searchers may elect, however, to enter a request term as a descriptor when it is mapped to a descriptor through a partial match [G], in which case it is usually mapped to a narrower descriptor. Searchers select this option because: (1) the term is added to the formulation to increase recall or to increase precision [GR1]; (2) the descriptor was spotted as an index term assigned to relevant articles [GD1]; (3) the searcher prefers to use descriptors and the selected one is the best match [GS1]; or (4) a combination of these reasons apply.

If suitable, however, searchers use a free-text key to inclusively search concepts that are not grouped together by the hierarchy of the controlled vocabulary [H]. This option is always selected to improve recall [HR1]. If, for example, the request term "students" is mapped to descriptors such as "foreign students," "college students," or "undergraduates," and the descriptor "student" does not exist, the free-text term can be used to retrieve information about almost any type of student.

It should be noted that in many search systems, use of the free-text key "student" also would retrieve citations that are indexed with descriptors which include the term. In other systems it is possible to retrieve only citations whose *indexing* includes this term. This is a source for constant confusion for searchers because the routine changes from one search system to another, and in one search system over a period of time.

(c) The descriptor is a broader term. When a single-meaning request

term is mapped to a broader descriptor, searchers may prefer to preserve the specificity of the request and use free-text search keys [I]. They do so when they want to increase precision [IR1], or when they subscribe to the axiom that the use of free-text terms increases recall [IS1].

A further concern with precision may lead searchers to enter free-text terms in combination (using the AND operator) with the broader descriptor to which it is mapped [J]. While precision is an important reason for the selection of this option [JR1], searchers may also use such a combination if they do not trust the indexing of the database [JD1].

Searchers, of course, may enter directly a broader descriptor [K]. Most often they select this option to increase recall [KR1]. Entering a broader descriptor is useful for recall enhancement in a variety of circumstances. Searchers may want to have an initial set that is broad because the request includes a relatively large number of concepts, or because the combination that is required by the request is limiting enough (if, say, the concepts are not likely to occur together). Another situation which calls for a broader descriptor is when an inclusive, rather than general, search is required to secure recall. For instance, when searchers enter the descriptor "students" for an inclusive search of a request about disabled students, they actually enter a broader term.

Depending on the terms, searchers may enter a broader descriptor to ensure both precision and recall [KR2]. Under such circumstances, the searchers perceive that the particular term would generate a set with low precision (if, for example, it is a single-word term). In addition, searchers may enter a broader descriptor when it is used only as a limiting factor [KR3].

The indexing in a particular database may also help searchers to select this option. They may enter a broader descriptor if it is found in the indexing of relevant citations [KD1], or because they generally prefer to use descriptors [KS1].

(d) Additional factors. A single-meaning term that is mapped to a descriptor--through any kind of match--provides searchers with more choices than those provided by terms which are not mapped to descriptors. Searchers, therefore, are free to consider other request-related factors.

If searchers, for example, think that a particular descriptor is assigned inconsistently by indexers, they may consider the use of free-text key to be more trustworthy [Z1]. Or, they may prefer to enter a descriptor when: a term has many synonyms [Z2]; a concept and its use is not clear to the searcher [Z3]; a concept is likely to be implied rather than explicitly mentioned in the searched text [Z4].

Recall and precision requirements can also be considered when the match between a term and descriptors presents no problems. Search keys can be used to increase recall in three ways: a searcher may add free-text synonyms to descriptors [Z5]; add the next broader descriptor in the hierarchy [Z6]; or use generic descriptors in an inclusive mode

[Z7].

Searchers elect to increase recall by adding free-text synonyms to a descriptor when they see the need to complement indexing [Z5R1]: They want to include citations that mention the concept, either in their titles or abstracts, even though the descriptor was not assigned to them. For some searchers this is the most straightforward approach to ensure recall: When a term--or a combination of terms--is specific, they require that the search key would occur in the descriptor, title, and the abstract fields. Searchers in the study selected this option at times because the user--who was present at the terminal--specifically insisted on using free-text terms as well as descriptors [Z5R2].

Database-related considerations may also lead searchers to the selection of this option. Searchers may decide to use free-text synonyms because they plan to search a number of databases [Z5D1] and wish to use the same query formulation across databases. Or, they may add free-text synonyms because they do not trust the indexing [Z5D2].

In contrast, adding the next broader descriptor in the hierarchy is selected as an option only when the searcher thinks that the user will be interested in the broader descriptor as well [Z6R1].

Use generic descriptors in an inclusive mode might be desirable for a number of reasons. When searchers create a set that they wish to combine with other sets in order to limit the scope of the retrieval, they may use a generic descriptor so the limiting set is not too restrictive [Z7R1].

Databases and their thesauri also play an important role in the choice of this option. A searcher who is interested in material about undergraduate students, for example, may want to secure high recall and retrieve all citations which are indexed under any descriptor which includes the term "students," whether or not the specific descriptor "undergraduate students" or the broader descriptor "students" exist [Z7D1]. Obviously, this is a specific use of the generic search: It can be carried out only for multi-words phrases and when a part of the phrase is generic by nature.

Inclusive searching might be induced by some databases which specifically recommend it and provide commands that perform such searching automatically. In these databases, a single command retrieves all the citations with descriptors that are narrower than the descriptor entered.

Searchers elect to increase precision by limiting the retrieval to descriptors only [Z8], or by limiting it to major descriptors [Z9]. The first option ensures that the articles whose citations are retrieved indeed deal with the subject matter, rather than merely mention it [Z8R1]. Alternately, limiting to a major descriptor is used to reduce the number of citations retrieved [Z9R1], or to make sure that a concept is central to the articles whose citations are retrieved [Z9R2].

Additional means to increase precision are to introduce role indicators [Z10], to specify document type [Z11], and to use free-text synonyms in a designated field in combination with descriptors [Z12]. The last option is considered by some to be a quick way to extract a subset that includes citations that are highly relevant from a relevant set [Z12R1]. For example, one may extract a highly relevant subset from the set retrieved with the descriptor "students" by adding the requirement that the term "students" appears in the titles of the articles as well. Searchers who do not trust the indexing of a particular database might choose this option [Z12D1].

3.1.3 A Single-Meaning Term That is Not Mapped to a Descriptor

The most direct approach is to enter a free-text search key when a term cannot be mapped to a descriptor [L]. Searchers, however, have some choices: They can enter a free-text term to probe indexing, or they can try and enter the term as a descriptor anyway. It is important, therefore, to examine the reasons for entering a free-text term directly without trying the other options.

A number of request-specific conditions may encourage a searcher to enter a free-text term directly. A searcher may do so if he believes that most specific retrieval is desired [LR1], or if the term itself is specific and well defined, that is, a term that is "ideal" for free-text searching [LR2]. This argument was frequently advanced by searchers who participated in the study when the term was a multi-word phrase and it was possible to use word-proximity operators.

A term that is not mapped to a descriptor can be added as a free-text term during the online session at the terminal. Searchers may decide to add such a term because it appears in titles or abstracts of relevant citations or because it is commonly used in the literature [LR3]. They may also add it only as a related term that is used to increase recall (e.g., names of particular examinations) [LR4]. In addition, searchers may enter free-text terms if the use of related descriptors results in a poor retrieval [LR5].

The nature of the controlled vocabulary for a database is also an important factor in the selection of free-text terms. A searcher may enter directly a free-text term rather than probe indexing because the term would not be a descriptor [LD1]. This would happen when: a thesaurus excludes a specific type of terms such as geographic names or other proper names; the concept belongs to a subject area that is not covered by the thesaurus; or the thesaurus is outdated and, therefore, would not include terms that represent "new" concepts. Further, searchers who do not trust the thesaurus' vocabulary or the indexing in a database may prefer to enter free-text terms directly [LD2].

Some searchers have adopted general rules that they apply whenever a term is not mapped to a descriptor. They may believe that: if a term represents a concept accurately there is no need to probe indexing [LS1]; free-text searching is best for high recall [LS2]; or terms that have been suggested by users can be entered as free-text term with no further probing [LS3].

Searchers who prefer to use descriptors, on the other hand, would enter free-text terms only to probe indexing, hoping to find descriptors that were assigned to relevant citations [M].

In some cases, searchers may use a free-text key to search for a single-meaning term that cannot be mapped to a descriptor in a particular way: They require that it occurs in a field other than the common ones, such as the Journal title field [N]. Suppose a user is interested only in the psychological aspects of students taking final examinations, and suppose that the term "psychology" cannot be mapped to a descriptor. Searchers may predict that searching for the occurrence of "psychology" in the text would retrieve a large number of irrelevant citations, and decide instead to retrieve citations to articles whose authors are affiliated with organizations which include the stem "psych" in their titles, or articles that were published in sources whose titles include this stem.

After unsuccessful attempts to find a descriptor, searchers may enter a term as a descriptor, even though it does not appear in the thesaurus [O]. They would choose this option either because they assume that the term might have been added to the thesaurus without their knowledge (for instance, before the supplements have been published) [OD1], or because the term is a descriptor in another database [OD2].

3.1.4 It is Not Known If a Term is Mapped to a Descriptor

When searchers elect not to check the thesaurus for a descriptor, they may: (a) enter free-text terms directly [P]; (b) use free-text terms to probe indexing [Q]; or (c) enter as a descriptor a term that might be a descriptor [R].

(a) Enter free-text terms directly. For some requests, searchers believe it is best to enter free-text terms without checking the thesaurus. They select this option when: they decide to enter the terms while they are online and have no time to examine the thesaurus [PR1]; the search is of the "quick-and-dirty" variety, or they are "just fishing" [PR2]; or the term is used to eliminate irrelevant citations--the term should not, therefore, appear in titles and abstracts of citations [PR3].

The availability of thesauri and their quality also lead searchers to enter a free-text term without looking for descriptors. Searchers would do so if: they do not trust the thesaurus and the indexing in a database [PD1]; they have decided to search a number of databases for one request--a decision they may make before or during the actual online session [PD2]; the thesaurus is not available to them [PD3]; or when they think that they are familiar with the thesaurus and are convinced that it would not have an adequate descriptor [PD4]. When they decide to change databases during the terminal session, searchers may enter a search statement that was constructed for the first database--including both descriptors and free-text terms--to be searched in the second database without checking its thesaurus [PD4].

Some searchers have general rules which favor searching with free-text terms only. They may prefer to use terms that have been suggested by the user because they believe that the use of these terms results in more relevant citations [PS1]. Or, they believe that free-text terms are better for recall [PS2].

(b) Use free-text terms to probe indexing. Searchers enter free-text terms to probe indexing because they are not sure which descriptor to use [QR1], because the thesaurus is not available to them [QD1], or because they generally prefer to start with free-text terms and only then check for descriptors [QS1].

(c) Enter as a descriptor a term that might be a descriptor. Searchers may enter a term as a descriptor when they add the term to the query formulation during the online session and they feel time is too precious to check the thesaurus [RR1]. They may resort to this option also when they perform a multi-database search [RR2].

Lastly, if terms are descriptors in another database [RD1], or if the thesaurus is not available [RD2], searchers will enter descriptors without checking the thesaurus, as they would do if they "know" that a term is a descriptor (or think it should be) [RD3].

3.2 Searchers' Selection of Search Keys

This section presents the frequency with which search keys, options, and reasons for options were selected. These descriptive statistics are based on data collected from 47 searchers performing a total of 281 searches. The data about reasons for option selection, however, were collected from 39 searchers performing a total of 201 searches.

3.2.1 Frequency of Search-Key Selection

Searchers selected a total of 3,635 search keys to perform the 281 searches. Of these, 1,607 (44% of all search keys selected) were descriptors, and 2,028 search keys were free-text terms.

Some of the databases searched, however, did not provide controlled vocabulary: 446 search keys were selected for databases that provide no choice in the selection of search keys. If we eliminate these search keys, the proportion between descriptor and free-text terms changes: Of the 3,189 search keys selected, 1,607 (50.40%) were descriptors, and 1,582 (49.60%) were free-text terms. That is:

> **Searchers did not display a general preference for one type of search keys: When they had a choice, they selected descriptors and free-text terms in the same frequency.**

In addition, searchers selected an average of 13.31 search keys per search, with a median of 9.20 and standard deviation of 12.80. The minimum average number of search keys per search for a searcher was 2.80 and the maximum was 68.75.

3.2.2 Frequency of Option Selection

The first four columns in Table 2 list the number of times each option was selected, the frequency with which each option was selected when all databases are considered, and the frequency with which it was selected in databases that have controlled vocabularies. These statistics show that the most frequent options were:

[F] use descriptors when a single-meaning term is mapped to a descriptor through an exact match (35.18%);

[P] use free-text terms when it is not known whether a single-meaning term is mapped to a descriptor (19.79%); and

[L] use free-text terms when a single-meaning term cannot be mapped to a descriptor (16.49%).

Table 2. Frequency of options and reasons

OPTION	NO.	% (1)	% (2)	CATEGORY	NO.	% (3)	REASON	NO.	% (4)
[A]	2	.05	.06						
[B]	6	.16	.18						
[C]	13	.36	.40	Request	3	23.08	CR1	3	100.00
				Database	10	76.92	CD1	10	100.00
[D]	1	.03	.03						
[E]	1	.03	.03						
[F]	1122	30.86	35.18						
[G]	44	1.21	1.37	Request	19	67.86	GR1	19	100.00
				Database	3	10.71	GD1	3	100.00
				Searcher	6	21.43	GS1	6	100.00
[H]	13	.35	.41	Request	13	100.00	HR1	13	100.00
[I]	3	.08	.09	Request	3	75.00	IR1	3	100.00
				Searcher	1	25.00	IS1	1	100.00
[J]	22	.60	.69	Request	1	50.00	JR1	1	100.00
				Database	1	50.00	JD1	1	100.00
[K]	96	2.64	3.10	Request	36	76.59	KR1	28	77.77
							KR2	6	16.66
							KR3	2	5.55
				Database	2	4.25	KD1	2	100.00
				Searcher	9	19.15	KS1	9	100.00
[L]	972	27.62	16.49	Request	187	49.08	LR1	84	44.92
							LR2	66	35.29
							LR3	27	14.44
							LR4	8	4.28
							LR5	2	1.07
				Database	105	27.56	LD1	99	94.28
							LD2	6	5.72
				Searcher	89	23.36	LS1	56	62.29
							LS2	17	19.10
							LS3	16	17.98
[M]	16	.44	.50						
[N]	1	.03	.03						
[O]	8	.22	.25	Database	4	100.00	OD1	3	75.00
							OD2	1	25.00
[P]	631	17.36	19.79	Request	101	12.58	PR1	62	61.39
							PR2	29	28.71
							PR3	10	9.99
				Database	461	57.41	PD1	129	27.98
							PD2	117	25.38
							PD3	108	23.43
							PD4	107	23.21
				Searcher	241	30.01	PS1	179	74.27
							PS2	62	25.73
[Q]	34	.93	1.07	Request	29	76.31	QR1	29	100.00
				Database	4	10.53	QD1	4	100.00
				Searcher	5	13.16	QS1	5	100.00
[R]	141	3.88	4.42	Request	14	9.10	RR1	13	92.86
							RR2	1	7.14
				Database	140	90.90	RD1	57	40.71
							RD2	42	30.00
							RD3	41	29.28
[Z1]	10	.27	.31						
[Z2]	1	.03	.03						
[Z3]	1	.03	.03						
[Z4]	1	.03	.03						
[Z5]	302	8.31	9.47	Request	119	72.12	Z5R1	117	98.32
							Z5R2	2	1.68
				Database	46	27.88	Z5D1	35	76.09
							Z5D2	11	23.91
[Z6]	6	.16	.19	Request	5	100.00	Z6R1	5	100.00
[Z7]	146	4.01	4.58	Request	6	10.34	Z7R1	6	100.00
				Database	52	89.66	Z7D1	52	100.00
[Z8]	1	.03	.03	Request	1	100.00	Z8R1	1	100.00
[Z9]	31	.85	.97	Request	11	100.00	Z9R1	6	54.54
							Z9R2	5	45.45
[Z10]	1	.03	.03						
[Z11]	1	.03	.03						
[Z12]	2	.05	.06	Request	2	66.67	Z12R1	2	100.00
				Database	1	33.33	Z12D1	1	100.00
[Z13]	14	38	.44						

(1) Percent of all search keys selected (47 searchers)
(2) Percent of search keys selected for databases with thesauri (47 searchers)
(3) Percent of category within the option (39 searchers)
(4) Percent of reasons within the category (39 searchers)

That is:

> **About 70% of the times, searchers selected the most straightforward options: If a term was mapped to a descriptor exactly, they entered a descriptor and if it could not be mapped, or when they did not consult a thesaurus, they entered free-text terms.**

Of particular interest is option [P] because it represents the instances where searchers decided to enter free-text terms without checking the thesaurus. It is useful, therefore, to spell out the reasons searchers cited to explain their decision not to have a choice in the selection of search keys.

Of the 803 instances in which searchers cited reasons for this option, 179 times (22.29% of the reasons for this option) they decided to avoid consulting a thesaurus because they held a general belief that entering the user's terms directly gives more relevant citations. While this belief was cited most frequently, 57.41% of the reasons given for this option were related to the databases searched: 129 times (16.06% of the reasons for this option) searchers claimed that they did not trust the thesaurus or the indexing; 117 times (14.57%) they said they did not consult a thesaurus because they were performing a multi-database search; 108 times (13.45%) they did not have the relevant thesaurus; and 107 times (13.32%) they did not think the term would be in the thesaurus. That is:

> **While the most frequent reason for not consulting a thesaurus was the belief that user's terms are best for relevant retrieval (22% of the reasons for this option), distrusting the thesaurus and the indexing (16%) and having to search several databases (15%) were also important reasons for searchers' decision to turn to this option.**

Among the options that are not straightforward, seven options were most prominent:

[Z5] add free-text synonyms to descriptors when a single-meaning term is mapped to a descriptor and recall needs to be improved (9.47%);

[Z7] use generic descriptors in an inclusive mode when a single-meaning term is mapped to a descriptor and recall needs to be improved (4.58%);

[R] enter as a descriptor a term that might be a descriptor when it is not known whether a single-meaning term is mapped to a descriptor (4.42%);

[K] use descriptors when a single-meaning term is mapped to a broader descriptor (3.01%);

[G] use descriptors when a single-meaning term is mapped to a descriptor through a partial match (1.37%);

[Q] use free-text terms to probe indexing when it is not known whether a single-meaning term is mapped to a descriptor (1.07%); and

[Z9] limit to major descriptors when a single-meaning term is mapped to a descriptor and precision needs to be improved (.97%).

The options that provide for high recall ([Z5], [Z7], and [K]), comprise 17.06% of the search keys selected. That is:

> **Among the options that are not straightforward, over a half were selected to enhance recall.**

Figure 2 is a network display of the Selection Routine that reflects the frequency with which each option was selected: solid lines for options that were selected more than 10% of the times, broken lines for those selected more than 1% of the time, and dotted lines for options that were selected less than 1% of the time. Figure 3 is the same display, including only the relatively frequent options--those that were selected more than 1% of the time.

3.2.3 Frequency of Reasons for Option Selection

The last six columns in Table 2 provide data about the reasons for selecting a certain option for conditions that result in more than one option. These data were derived from 37 searchers. The first of these six columns lists the category which represents the reason: whether the reason was related to a request, the database, or a searcher. The second column tallies the total number of times that reasons in a particular category were given for the option. The third column includes the percentage of each category within the option. The next column lists the code of each individual reason, followed in the next column by the number of times the reason was mentioned. The last column represents the percentage of each reason within its category.

It should be noted that the total number of reasons used for a particular option was frequently different from the number of times the option was selected. There are two sources for this discrepancy. First, the data about the number of times an option was selected was derived from observing 47 searchers, while the data about reasons were collected from 37 searchers. Thus, for example, while option [G] was selected 44 times, the total number of reasons is only 28 (19+3+6). Second, a selection of an option may be caused by more than one reason. A searcher may decide, for example, to select a free-text term because she believes that free-text terms increase recall (a searcher-related reason), but also because she does not trust the indexing (a database-related reason). Thus, option [I], for instance, has only three instances, but four (3+1) reasons.

A summary of the reasons used for the selection of search keys shows that of a total of 1733 reasons, 553 (31.91%) were related to the request, 829 (47.84%) were database-related, and 351 (20.25%) were searcher-related. That is:

> **When searchers had options in the selection of search keys, their choice was most frequently (48% of the time) determined by the databases they were searching and least frequently (20%) by their habitual searching behavior.**

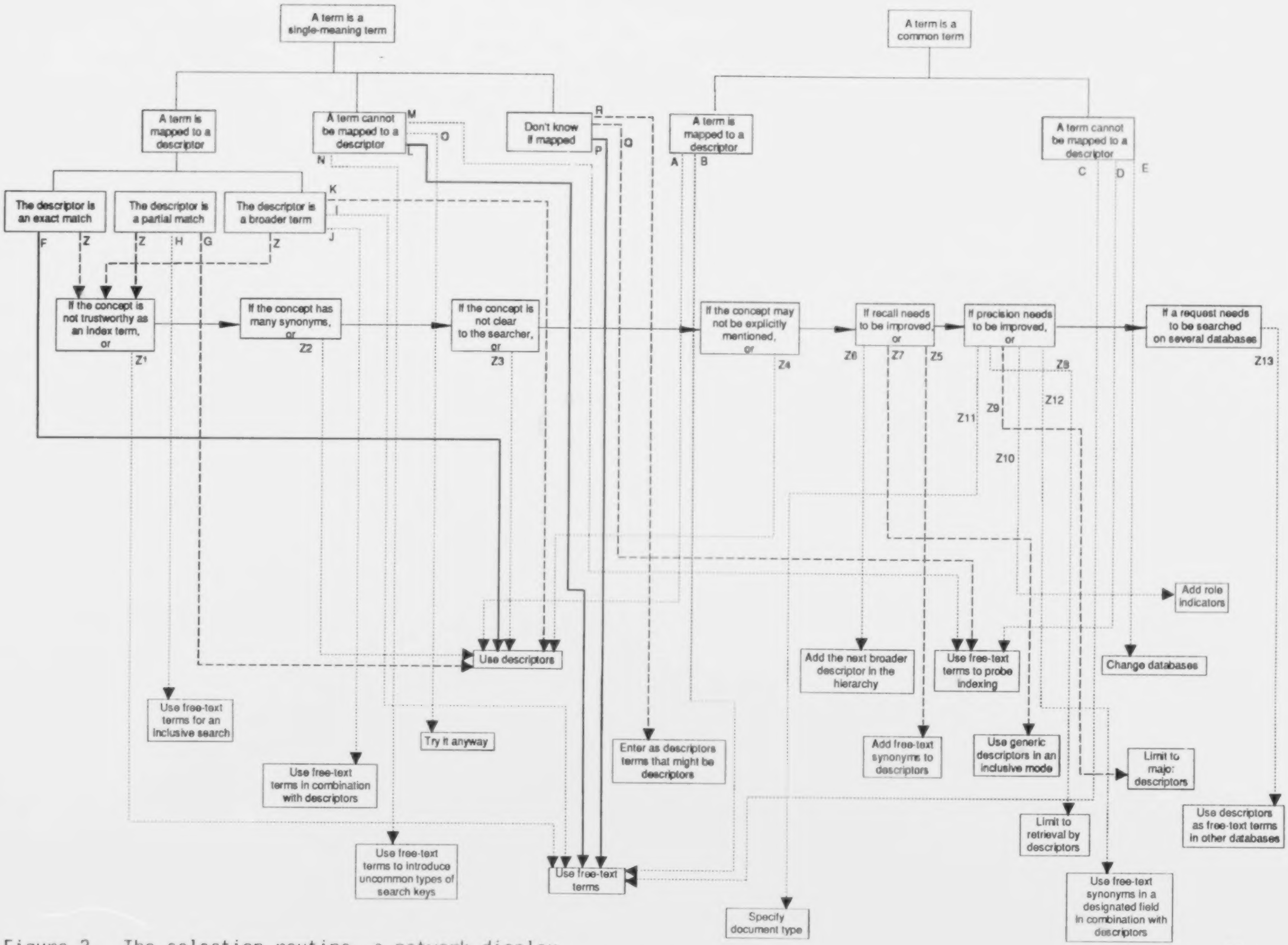
A term is a single-meaning term
A term is a common term
A term is mapped to a descriptor
A term cannot be mapped to a descriptor
Don't know if mapped
A term is mapped to a descriptor
A term cannot be mapped to a descriptor
The descriptor is an exact match
The descriptor is a partial match
The descriptor is a broader term
If the concept is not trustworthy as an index term, or
If the concept has many synonyms, or
If the concept is not clear to the searcher, or
If the concept may not be explicitly mentioned, or
If recall needs to be improved, or
If precision needs to be improved, or
If a request needs to be searched on several databases
Add role indicators
Use descriptors
Add the next broader descriptor in the hierarchy
Use free-text terms to probe indexing
Change databases
Use free-text terms for an inclusive search
Try it anyway
Enter as descriptors terms that might be descriptors
Add free-text synonyms to descriptors
Use generic descriptors in an inclusive mode
Limit to major descriptors
Use free-text terms in combination with descriptors
Limit to retrieval by descriptors
Use descriptors as free-text terms in other databases
Use free-text terms to introduce uncommon types of search keys
Use free-text terms
Use free-text synonyms in a designated field in combination with descriptors
Specify document type
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
Z
Z
Z
Z
Z1
Z2
Z3
Z4
Z5
Z6
Z7
Z8
Z9
Z10
Z11
Z12
Z13

Figure 2. The selection routine--a network display

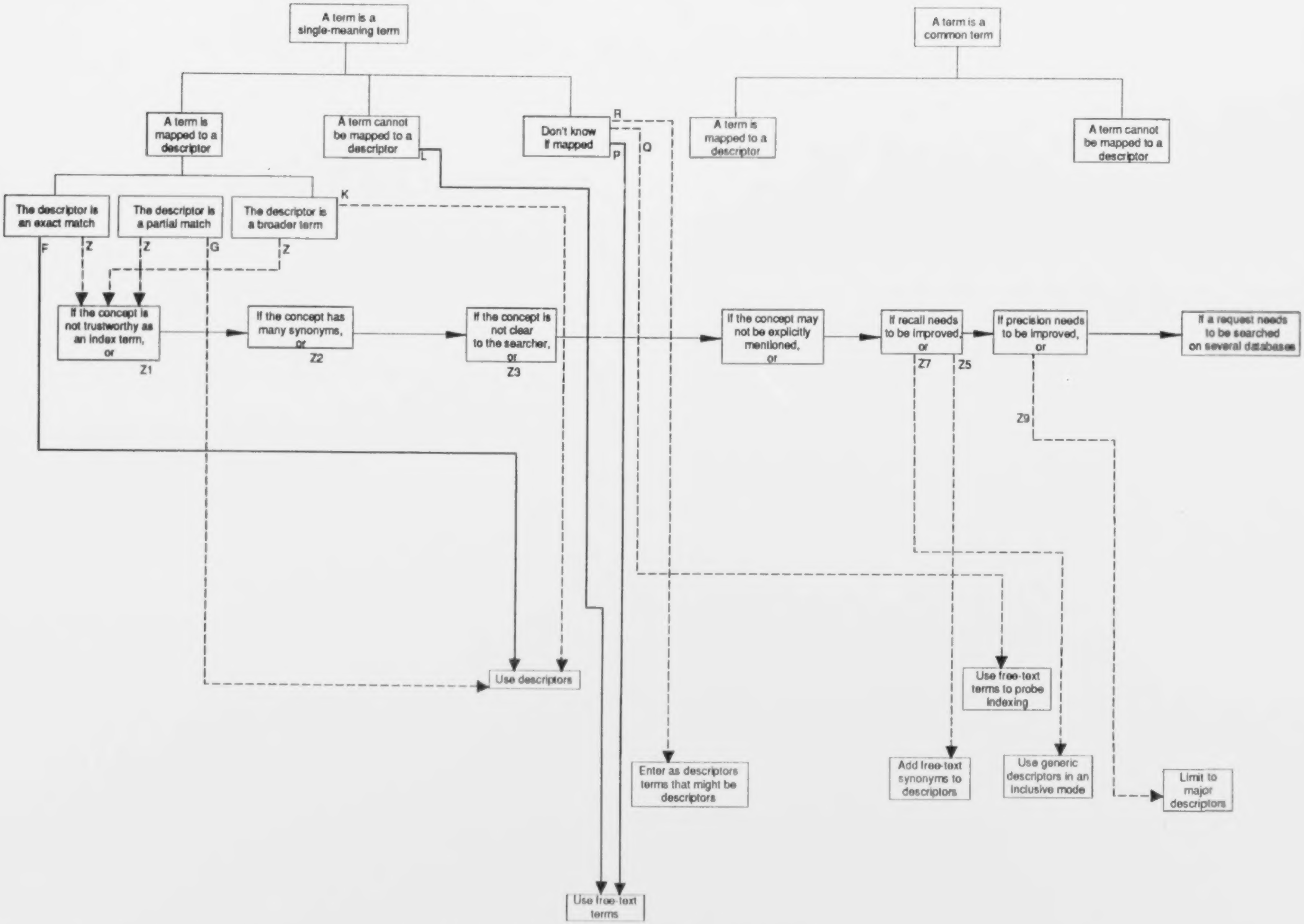
A term is a single-meaning term
A term is a common term
A term is mapped to a descriptor
A term cannot be mapped to a descriptor
Don't know if mapped
R
Q
P
L
A term is mapped to a descriptor
A term cannot be mapped to a descriptor
The descriptor is an exact match
The descriptor is a partial match
The descriptor is a broader term
K
F
Z
Z
G
Z
If the concept is not trustworthy as an index term, or
Z1
If the concept has many synonyms, or
Z2
If the concept is not clear to the searcher, or
Z3
If the concept may not be explicitly mentioned, or
If recall needs to be improved, or
Z7
Z5
If precision needs to be improved, or
Z9
If a request needs to be searched on several databases
Use descriptors
Use free-text terms to probe indexing
Enter as descriptors terms that might be descriptors
Add free-text synonyms to descriptors
Use generic descriptors in an inclusive mode
Limit to major descriptors
Use free-text terms

Figure 3. Frequent options in the selection routine--a network display

In the description of the Selection Routine, each reason is described under the option, which in turn is delineated with the condition that generates the option. The same reason, however, may lead to different options, depending on the specific conditions. It is useful, therefore, to examine each reason, the conditions, the options to which it may lead, and the frequency in which the 37 searchers selected each reason-option combination.

(a) Request-related reasons. Among the reasons that were used by searchers to explain the selection of search keys, 16 related to attributes of individual requests. These are the reasons as presented by the searchers, and the resulting options they selected:

(1) "Recall needs to be improved" (31.28% of request-related reasons), induced four options: add free-text synonyms to descriptors when a single-meaning term is mapped to a descriptor ([Z5R1]--117 times); use descriptors when a single-meaning term is mapped to a broader descriptor ([KR1] and [KR2]--34 times); use free-text terms for an inclusive search when a single-meaning term is mapped to a descriptor through a partial match ([HR1]--14 times); and use free-text terms for a single-meaning term that cannot be mapped to a descriptor ([LR4]--8 times).

(2) "Most specific retrieval is desired" (15.19%) was a reason to use free-text terms to represent a single-meaning term that could not be mapped to a descriptor ([LR1]--84 times).

(3) "The term was added while online" (13.56%) was a reason to enter free-text keys when a thesaurus was not consulted for a single-meaning term ([PR1]--62 times), and to enter free-text terms to probe indexing under the same condition ([RR1]--13 times).

(4) "The term is specific and well-defined" (11.93%) was a reason to enter free-text terms to represent a single-meaning term that could not be mapped to a descriptor ([LR2]--66 times).

(5) "I am not sure what descriptors to use" (5.24%) caused searchers to enter free-text terms without consulting the thesaurus to probe indexing for a single-meaning term ([QR1]--29 times).

(6) "I don't have time to look for descriptors--I am just fishing" (5.24%) was a reason to avoid consulting the thesaurus for a single-meaning term and to enter free-text terms ([PR2]--29 times).

(7) "The term appeared in titles and abstracts of relevant articles" (4.88%) was used to explain entering free-text terms to represent a single-meaning term that was not mapped to a descriptor ([LR3]--27 times).

(8) "The term is only added to the formulation to increase recall" (3.43%) facilitated the use of descriptors when a single-meaning term was mapped to a descriptor through partial match ([GR1]--19 times).

(9) "Precision needs to be improved" (2.35%) was used as a reason to explain three of the options that resulted when a single-meaning term was mapped to a broader descriptor: use free-text terms ([IR1]--6 times); use descriptors ([KR2]--6 times); and use free-text terms in combination with descriptors ([JR1]--1 time).

(10) "The term is used to eliminate irrelevant citations" (1.81%) was a reason to use free-text terms to represent a single-meaning term without consulting the thesaurus ([PR3]--10 times).

(11) "The term is used as a limiting factor" (1.63%) was a reason to enter a common term that was mapped to a descriptor as a free-text term ([BR1]--7 times), and to use descriptors when a single-meaning term was mapped to a broader descriptor ([KR3]--2 times).

(12) "The query formulation includes a relatively large number of components" (1.63%) was used to explain both the use of generic descriptors in an inclusive mode for a single-meaning term that was mapped to a descriptor and when recall needed to be improved ([Z7R1]--6 times), and the use of free-text terms to represent a common term that was not mapped to a descriptor ([CR1]--3 times).

(13) "The size of the set needs to be reduced" (1.08%) caused searchers to limit a descriptor that represented a single-meaning term to a major descriptor ([Z9R1]--6 times)

(14) "To make sure that a concept is central to articles" (.90%) was cited as a reason to limit to major descriptors retrieval for single-meaning terms that were mapped to descriptors ([Z9R2]--5 times).

(15) "Had gotten poor retrieval using related descriptors" (.36%) caused searchers to use free-text terms to represent a single-meaning term that could not be mapped to a descriptor ([LR5]--2 times).

(16) "User insisted on using the terms" (.36%) led searchers to add free-text synonyms to descriptors that represented single-meaning terms ([Z5R2]--2 times).

The request-related reason that was used most frequently (the first reason), as well as the eighth reason, were used to increase recall. Therefore:

> **Among the request-related reasons, the need to enhance recall was the most frequent reason (35% of request-related reasons) for the selection of a certain option.**

(b) Database-related reasons. Nine reasons given by searchers to explain their selection of search keys related to attributes of the databases searched.

(1) "A term would not be in the thesaurus" (24.85% of database-related reasons) led searchers to enter free-text terms without consulting a thesaurus ([PD4]--107 times), and when they could not map a single-

meaning term to a descriptor ([LD1]--99 times).

(2) "Needed to perform a multi-database search" (19.54%) caused searchers to: use free-text terms to represent a single-meaning term without checking a thesaurus ([PD2]--117 times); to add free-text synonyms to descriptors of single-meaning terms when recall needed to be improved ([Z5D1]--35 times); and to enter a common term that was not mapped to a descriptor as a free-text term ([CD1]--10 times).

(3) "The thesaurus is not available" (18.58%) was cited as a reason for not consulting the thesaurus of a database, which in turn led to the options: use free-text terms to represent a single-meaning term ([PD3]--108 times); enter as a descriptor a single-meaning term that might be a descriptor ([RD2]--42 times); and use free-text terms to probe indexing ([QD1]--4 times).

(4) "I don't trust the descriptors and/or the indexing" (17.85%) generated a number of options: use free-text terms to represent a single-meaning term without consulting a thesaurus ([PD1]--129 times); add free-text synonyms to descriptors when recall needs to be improved ([Z5D2]--11 times); use free-text terms to represent a single-meaning term that cannot be mapped to a descriptor ([LD2]--6 times); use free-text terms in combination with descriptors to represent a single-meaning term that is mapped to a broader descriptor ([JD1]--1 time); and use free-text synonyms in a designated field in combination with descriptors to increase precision ([Z12D1]--1 time).

(5) "The term is a descriptor in another database" (7.00%) was a reason to enter as a descriptor a single-meaning term that might be a descriptor without consulting a thesaurus ([RD1]--51 times), and to enter a single-meaning term as a descriptor even though it could not be mapped to a descriptor ([OD2]--1 time).

(6) "Wanted to include all descriptors which contain a certain phrase" (6.27%) was used as a reason to enter generic descriptors in an inclusive mode when recall needed to be improved ([Z7D1]--52 times).

(7) "I 'know' the terms are descriptors" (4.94%) was the reason for entering a single-meaning term as a descriptor without consulting a thesaurus ([RD3]--41 times).

(8) "A term was found as an index term in relevant articles" (.60%) caused searchers to enter a descriptor with partial match ([GD1]--3 times), and to enter a broader descriptor ([KD1]--2 times).

(9) "A term might have been added to the thesaurus" (.36%) led searchers to enter as a descriptor a single-meaning term that could not be mapped to a descriptor ([OD1]--3 times).

(c) Searcher-related reasons. Five reasons given by searchers explaining their selection of search keys were actually general rules or beliefs held by the individual searchers.

(1) "Terms suggested by users are the best for retrieval" (55.55% of searcher-related reasons) explained the use of free-text terms without consulting a thesaurus ([PS1]--179 times), and the use of free-text terms when a single-meaning term was mapped to a broader descriptor ([LS3]--16 times).

(2) "The use of free-text terms increases recall" (22.79%) justified entering free-text terms without consulting a thesaurus ([PS1]--62 times), when a single-meaning term could not be mapped to a descriptor ([LS2]--17 times), and when a term was mapped to a broader descriptor ([IS1]--1 time).

(3) "If a term represents a concept accurately and it is not mapped to a descriptor, there is no need to probe indexing" (15.95%) led searchers to enter free-text terms whenever a single-meaning term was not mapped to a descriptor ([LS1]--56 times).

(4) "I prefer to use descriptors" (4.27%) caused searchers to enter descriptors when a single-meaning term was mapped to a broader descriptor ([KS1]--9 times), and when it was mapped through partial match ([GS1]--6 times).

(5) "I prefer to start with free-text terms and then check descriptors" (1.42%) explained why searchers entered free-text terms to probe indexing without consulting a thesaurus first ([QS1]--5 times).

The number of reasons in a category reflects the variability that is introduced to online searching by the category. A category that includes a small number of reasons introduces a relatively small variability because the reasons can be easily predicted, and vice versa. Therefore:

> **Search requests introduced the largest variability to the search process and beliefs held by individual searchers introduced the smallest variability.**

3.2.4 Frequency of Search-Key Selection for Databases

A total of 70 databases were searched by the 47 searchers. Five databases did not have controlled vocabulary, and 31 were searched infrequently (i.e., with less than 10 search keys). Statistics about search-key selection in the remaining 34 databases is provided in Table 3. This table includes the following information for each database: the total number of search keys selected; the percentage of descriptors selected; and the percentage of free-text terms that were entered directly without consulting a thesaurus.

Table 3. Search-Key Selection for Databases

Database	Total # of search keys	Descriptors (%)	Selected with no thesaurus (%)
ERIC	347	77.80	8.64
NTIS	119	17.65	13.44
COMPENDEX	59	30.51	40.68
AGRICULA	65	3.08	46.15
PSYCINFO	213	68.07	17.37
INSPEC	46	43.48	28.26
ABI/INFORM	124	44.35	17.74
PROMPT	40	27.50	50.00
SOCIOLOGICAL ABSTRACTS	86	47.67	23.25
AMERICA: HISTORY & LIFE	21	33.33	66.66
HISTORICAL ABSTRACTS	26	76.92	23.08
ASFA	34	11.76	17.65
MAGAZINE INDEX	72	19.44	62.50
PAIS	37	78.38	21.62
CAB ABSTRACTS	33	9.09	90.91
FOOD SCIENCE TECHNOLOGY	15	0.00	60.00
BIOSIS	195	35.38	7.18
LISA	16	12.50	12.50
CHILD ABUSE & NEGLECT	21	85.71	14.28
MLA BIBLIOGRAPHY	28	35.71	28.57
MANAGEMENT CONTENTS	52	50.00	36.53
GEOREF	56	23.21	58.93
US POLITICAL SCIENCE	38	65.79	31.58
AEROSPACE ONLINE	17	0.00	29.41
NATIONAL NEWSPAPER INDEX	25	68.00	20.00
WATER RESOURCES ABSTRACTS	21	0.00	14.28
LEGAL RESOURCE INDEX	12	25.00	0.00
HEALTH PLANNING & ADMINISTRATION	53	60.37	3.77
MEDLINE	674	67.50	1.93
RELIGION INDEX	21	61.90	19.04
NEWSEARCH	27	33.33	29.63
THE COMPUTER DATABASE	70	34.28	27.14
CA SEARCH	124	29.03	20.16
NASA	21	57.14	0.00

Table 3 provides information about individual databases. Because no concerted effort was made here to represent each database equally, the data collected in this study can hardly be used to generate general statements about individual databases. These data, however, can point to the frequency in which searchers use the controlled vocabulary of databases. Using these data, it was found that for the databases used in this study, the percentage of descriptors selected and the percentage of free-text terms entered without consulting a thesaurus are inversely related with each other, $r(32) = -.435$, $p < .01$. Therefore, it is plausible to suggest the hypothesis that:

> **Searchers are less likely to enter free-text terms without consulting a thesaurus when they search databases for which they usually use descriptors than when they search databases for which they use descriptors infrequently.**

If proven valid, this association will show that databases acquire a "reputation" among searchers: Some are typically searched with descriptors and for the others--those that are searched most commonly with free-text keys--searchers often do not bother to check the thesaurus.

3.3 Searchers' Selection of Moves

Table 4 is a list of the moves in online searching--i.e., modifications in search strategies. The Table includes 20 operational moves--moves that do not change the meaning of a request. These are divided into moves to reduce the size of a set (12 moves), moves to enlarge the size of a set (7 moves), and those to increase both precision and recall (1 move). Thirteen moves are conceptual moves, that is, moves that change the meaning of the request. Of these, five are moves to reduce the size of a set, six to enlarge the size of a set, and two moves to increase both precision and recall. A full explanation of the moves is available elsewhere [Fidel, 1985].

3.3.1 The frequency of Moves Selection

The 47 searchers made a total of 1,244 moves in their searches. Of these, 497 (39.95%) were conceptual moves and 747 moves (60.05%) were operational moves.

One operational move, however, is actually determined by the availability of databases, rather than by request consideration: the move to add a database (Add 5). This move is often imposed by the search system when a complete run of a database is split into a number of databases, each covering a different period of time. Searchers selected this move 312 times. If we eliminate this move from the list, the proportion between conceptual and operational moves changes: 435 moves (46.47% of the moves) were operational moves and 497 (53.33%) were conceptual moves. That is:

> **Searchers, in general, did not prefer one type of move on the other: About half of the moves they selected were conceptual moves and the other half--operational.**

While searchers in general do not prefer one type of move above the other, individual searchers may have a preference for one type of move. Of the 47 searchers, 25 selected moves of a particular type more than 70% of the time, and 35 searchers selected moves of a particular type more than 60% of the time. Three searchers selected moves of one type only.

Table 5 reports the frequencies in which searchers selected moves. For each move the number of times it was used and the frequency with which it was used in relation to the total number of moves are given.

Table 4. Moves in online searching

Operational moves		*Conceptual moves*	
	Moves to reduce the size of a set		
Weight 1	Limit a descriptor to be a major descriptor.	Intersect 1	Intersect a set with a set representing another query component.
Weight 2	Intersect free-text set with a broader descriptor.	Narrow 1	Intersect a descriptor set with a set created by more specific free-text terms.
Weight 3	Limit free-text terms to occur in a predetermined field.	Narrow 2	Qualify descriptors with role indicators.
Weight 4	Require that free-text terms occur closer to one another in the searched text.	Narrow 3	Select a narrower concept.
Weight 5	Limit to documents of a certain form.	Intersect 2	Intersect sets with role indicators.
Negate	Eliminate unwanted elements by using the AND NOT operator.		
Eliminate	Eliminate a term from the formulation.		
Limit 1	Limit to documents written in a particular language.		
Limit 2	Limit to documents published, or indexed, in a particular period of time.		
Limit 3	Limit to documents retrieved from a specific portion of the database.		
Limit 4	Limit to sources that have, or do not have, a certain term in their titles.		
Cut	Submit only part of the retrieved answer set, arbitrarily selected.		
	Moves to enlarge the size of a set		
Add 1	Add synonyms and variant spellings.	Expand 1	Enter a broader descriptor or term.
Add 2	Add descriptors as free-text terms.	Expand 2	Group together search terms to broaden the meaning of a set.
Add 3	Add terms occurring in records of relevant citations retrieved.	Expand 3	Group together a descriptor with an equivalent role indicator.
Add 4	Add terms from database's index that have a high number of postings.	Expand 4	Represent a query component explicitly only by qualifying another component with role indicators.
Add 5	Move to a new database.	Exclude	Exclude from a formulation concepts present in most documents in a database.
Include	Group together a descriptor with all the descriptors that are its narrower terms.	Expand 5	Supplement a specific answer set with sets representing broader concepts.
Cancel	Eliminate restrictions previously imposed.		
	Moves to increase both precision and recall		
Refine	Find a "better" search key.	Probe 1	Construct an indexing-probe set.
		Probe 2	Use the difference among the number of postings for a search term in various databases to decide how to represent components in each database.

Table 5. Frequency of move selection

Move	Number of times	%	Move	Number of times	%	Total
			Moves to reduce the size of a set			
Weight 1	35	2.81	Intersect 1	75	6.03	
Weight 2	2	.16	Narrow 1	8	.64	
Weight 3	37	2.97	Narrow 2	6	.48	
Weight 4	18	1.45	Narrow 3	52	4.18	
Weight 5	15	1.20	Intersect 2	4	.32	
Negate	34	2.73				
Eliminate	5	.40				
Limit 1	13	1.04				
Limit 2	63	5.06				
Limit 3	10	.80				
Limit 4	1	.08				
Cut	22	1.77				
SUBTOTAL	255	20.50		145	11.65	400
			Moves to enlarge the size of a set			
Add 1	63	5.06	Expand 1	56	4.50	
Add 2	7	.56	Expand 2	81	6.51	
Add 3	43	3.46	Expand 3	2	.16	
Add 4	1	.08	Expand 4	21	1.69	
Add 5	312	25.08	Exclude	4	.32	
Include	15	1.21	Expand 5	145	11.66	
Cancel	27	2.17				
SUBTOTAL	468	37.62		309	24.84	777
			Moves to increase both precision and recall			
Refine	25	2.01	Probe 1	37	2.97	
			Probe 2	7	.56	
SUBTOTAL	25	2.01		44	3.54	69
TOTAL	748	60.05		498	39.95	1246

Data in Table 5 show that 62.46% of the moves were employed to increase the recall of a retrieved set, while only 32.15% of the moves were directed at reducing the size of a set. Although the move Add 5 is often imposed by the distribution of information among databases, it is always made to improve recall and therefore should be counted here, even though it was eliminated from the comparison between the frequency with which conceptual and operational moves are made. Therefore:

The number of moves to increase recall was almost double the number of moves to increase precision.

Table 5 also shows that moves to reduce the size of a set are more often operational than conceptual. This relation suggests the hypothesis:

Searchers who prefer to make operational moves are more likely to employ moves to reduce the size of a set than searchers who prefer to make conceptual moves.

It is useful to note the moves that are most "popular" among searchers. Among the moves to reduce the size of a set, the move Limit 2 (limit to documents published, or indexed, in a particular period of time) is the most frequent one among the operational moves, and Intersect 1 (intersect a set with a set representing another query component) is most frequent among conceptual moves. The most popular move to increase the size of a set among the operational moves is Add 5 (move to a new database), and Expand 5 (supplement a specific answer set with sets representing broader concepts) is the most frequent among the conceptual moves.

The array of moves selected by each searcher was rather limited. Of the 43 moves available to searchers, the average number of moves that constituted a searcher's repertoire was 8.32 with median of 8.00 and standard deviation of 3.52. The maximum number of individual moves that one searcher employed was 17, and the minimum was three. That is:

On the average, each searcher employed less than 20% of the moves that were available to her or him.

4. FACTORS AFFECTING THE SELECTION OF SEARCH KEYS

The reasons provided by searchers when they explained their selection of search keys reflect searchers' perceptions. These perceptions are highly relevant because they guide searchers in their selection of search keys. Because they are subjective, however, these perceptions cannot be used as the sole source of evidence for determining the factors that affect searching behavior; they need to be supported by objective measurements.

To substantiate searchers' perceptions, statistical associations among eleven variables were measured. The analysis was based on 281 searches performed by 47 searchers. Most associations were analyzed on two levels: (1) the search level, where each search was considered a distinct instance (a total of 281 instances); and (2) the person level, where the values for each person were aggregated so that each person was considered a distinct instance (a total of 47 instances). One should note, however, that the instances on the search level are not independent because each set of five searches was performed by the same person.

The variables examined for this study were:

1. *The number of search keys.* Search level: the number of keys selected for a search. Person level: the average number of search keys selected by a searcher per search.

2. *Search-keys ratio.* The percentage of free-text terms selected. Search level: the number of free-text search keys, divided by the total number of search keys selected for a search. Person level: the total number of free-text search keys, divided by the total number of search keys selected by a searcher.

3. *Thesaurus look-ups.* The percent of terms entered *without* consulting a thesaurus. Search level: the number of free-text search keys entered during a search without consulting a thesaurus, divided by the number of search keys selected for the search. Person level: the total number of free-text terms entered by a searcher without consulting a thesaurus, divided by the total number of search keys entered by the searcher.

4. *The number of databases.* Search level: the total number of databases used for a search, determined by the number of times the move Add 5 (move to a new database) occurred in the search. Person level: the average number of databases used per search, determined by the average number of times a searcher made the move Add 5 per search.

5. *The number of moves.* Search level: the total number of moves made during a search. Person level: the average number of moves made by a searcher per search.

6. *Moves ratio.* The percentage of operational moves. Search level: the number of operational moves, divided by the total number of moves made during a search. Person level: the total number of operational moves, divided by the total number of moves made by a searcher.

7. *Precision moves.* The number of moves made to reduce the size of a set in a search (search level only).

8. *Recall moves.* The number of moves made to increase the size of a set in a search (search level only).

9. *Recall tendency.* The percentage of moves made to increase the size of a set. Search level: the number of recall moves, divided by the number of moves made during a search. Person level: the total number of recall moves, divided by the total number of moves made by a searcher.

10. *Subject area.* The subject area in which a searcher specializes (person level only). This variable used four categories: medicine, sciences, social sciences (for the social sciences and the humanities), and general (for searchers who habitually search requests in a variety of subjects, as is often the case in public libraries or for independent consultants).

11. *Environment.* The environment in which a searcher works (person level only). Two categories were intuitively defined: practical environments and theoretical ones. A practical environment is a working place where searchers are usually called to search requests that result from immediate and practical problems, as are most small or medium-size consulting companies or industries. In contrast, theoretical environments are establishments whose users are most often involved in research or investigation projects, as are universities or regulatory agencies. Some search environments could not be assigned any of the two categories; these were called general environments.

4.1 The Number of Search Keys

This variable, which measures the total number of search keys for a search and the average number of search keys selected by a searcher per search, is associated with: the number of moves, the environment, and the number of databases.

4.1.1 The Number of Moves

The number of search keys is directly correlated with the number of moves (search level: $r(279) = .684$, $p < .01$; person level: $r(45) = .777$, $p < .01$). One possible explanation for this correlation is that moves are made with search keys, which makes the association trivial. This notion, however, in not grounded in actual searching: Examination of the list of moves (Table 4) shows that among the 42 moves only 12 require the use of search keys for their execution. Therefore, this association points to a significant pattern in online searching behavior.

The number of moves during a search--or the average number of moves a searcher made per search--reflects the degree of interaction: the larger the number of moves, the more interactive a search--or a searcher--is. Therefore, this association shows that:

> **Interactive searches are likely to require a larger number of search keys than less interactive searches. Similarly, interactive searchers are likely to use a larger number of search keys than searchers who are less interactive.**

The number of search keys is also correlated directly with the number of precision moves in a search ($r(279) = .377$, $p < .01$), and with the number of recall moves ($r(279) = .602$, $p < .01$). This association is predictable, however, because we already know that the number of moves, whether they are precision or recall moves, is associated with the number of search keys. Yet, with coefficient of determination $r^2 = .142$ for precision moves, and $r^2 = .362$ for recall moves, the association with precision moves explains 14% of the instances, while that with recall moves explains 36%. Therefore, while the total number of moves is associated with the number of search keys, recall moves contribute to this association 2.55 times more than precision moves do. Though it is tempting to claim that this difference proves the commonly-held assumption that recall moves require more search keys than precision moves, one has to remember that the total number of recall moves recorded for the study population was double the number of precision moves (section 3.3.1). This feature alone can explain why recall moves contribute the larger part to the association between the number of moves and the number of search keys.

4.1.2 Environment

The number of search keys is associated with the environment in which a searcher works ($F(2, 44) = 5.22$, $p < .01$). Searchers who work in practical environments use an average of 6.76 search keys per search, those in theoretical environments use an average of 18.56 search keys per search, and those who work in general environments use an average of 11.76 search keys. A post-hoc test shows a significant difference between the practical and theoretical environments. Although the variable environment lacks a rigorous definition, this association shows that:

> **Searchers who are used to answering practical questions use a considerably smaller number of search keys per search than do searchers who habitually answer theoretical requests.**

This conclusion is not surprising. It is commonly assumed that theoretical requests usually require high recall, and that high-recall requests require a relatively extensive use of search keys. Thus, even though these assumptions have not been substantiated before, this finding agrees with common knowledge. Further, this association suggests that:

> **The type of a request, whether practical or theoretical, may determine the number of search keys used.**

4.1.3 The Number of Databases

The number of search keys is directly associated with the number of databases used (search level: $r(279) = .324$, $p < .01$; person level: $r(45) = .464$, $p < .01$). This correlation, however, was partially induced by the method used in this study to analyze search protocols. In this analysis, we considered every entry of a search key as an instance of a search-key selection, whether or not the search key had been entered before. Thus, if a searcher entered the same search key in, say, five databases, the search key was counted five times. (There was one exception, though: We did not count search keys when a search in one database was saved and then automatically transferred to the next database without re-entering the query formulation.) Following this analysis, the use of each new database automatically increased the number of search keys counted.

This association is, therefore, trivial. In fact, our observations of actual searches led us to believe that the number of search keys may even relate inversely to the number of databases, because searchers at times added databases instead of adding terms when they wanted to enlarge the size of a set, and vice versa. Unfortunately, the method of data analysis used in this study prevents us from testing the validity of this notion.

We turn now to examine the variables that are not associated with the number of search keys.

4.1.4 Search-Keys Ratio

The percent of free-text terms selection does not significantly correlate with the number of search keys (search level: $r(279) = -.016$, NS; person level: $r(45) = -.166$, NS). This association leads to the conclusion that:

Searchers who prefer to use free-text terms and those who prefer descriptors use, on the average, the same number of search keys.

This finding agrees with the finding that the searchers in this study selected almost an equal number of descriptors and free-text terms (section 3.2.1). But it contradicts the well-known assumption that when searchers use free-text terms they are likely to use more terms than when they use descriptors, because with free-text terms they are free to choose any term that seems relevant to them. While this is a sound assumption, it is not supported by the data collected in this study. This result shows, then, that one of the assumed advantages of free-text searching does not hold in real-life searching.

Further, this finding highlights the essential role of controlled vocabularies and of indexing. One of the central purposes of vocabulary control is to control for synonyms. Thus, instead of searchers having to exercise terminological control while searching by thinking up all relevant synonyms for a concept, control is conducted at the design stage and each concept is represented with one term only. The finding that searchers who prefer the use of free-text terms enter, on the average, the same number of search keys as searchers who prefer descriptors may be explained, therefore, by the idea that searchers who prefer free-text terms do not exercise terminological control when they enter free-text terms because if they did, the average number of search keys they use would have increased.

While it is easy to conclude that searchers should perform their searches more thoroughly, this notion warrants the attention of designers of database and of expert systems. If searchers do not exercise terminological control in searching (and whether they shy away from it because they feel inhibited or because it takes a special talent to do so while searching under cost constraints, is immaterial), database designers should encourage the use of thesauri by designing reliable thesauri that are easy to use, and intermediary expert systems should be designed to help searchers in terminological control.

4.1.5 Moves Ratio

The percent of operational moves does not significantly correlate with the number of search keys (search level: r(279) = .033, NS; person level: r(45) = .220, NS). This shows that:

On the average, operationalist and conceptualist searchers are likely to use the same number of search keys.

4.1.6 Recall Tendency

The percent of recall moves does not significantly correlate with the number of search keys (search level: r(279) = .086, NS; person level: r(45) = .082, NS). On the surface, one would expect these variables to correlate directly to one another because it is commonly assumed that high-recall requests require a relatively large number of search keys. A more careful examination shows, however, that our finding does not completely contradict this assumption.

The lack of association between searches with a relatively large number of recall moves (or searchers who make, on the average, a high percent of recall moves) and the number of search keys may be explained by the observation that searchers increase recall _either_ by using more search keys, _or_ by making moves to increase recall. Since not all moves to increase recall require the use of additional search keys, recall can be improved without an increased use of search keys. For example, the move Expand 5 (supplement a specific answer set with sets representing broader concepts) is a conceptual move to increase the size of a set that was made 19% of the times recall moves were made (31% of recall moves, ignoring Add 5), and it does not require entering additional search keys. Therefore, this finding merely indicates that _when_ searchers make moves to increase recall they do not necessarily use additional search keys. This conclusion agrees with previous data. Therefore:

> **Searchers who frequently make recall moves do _not_ use a larger than average number of search keys.**

4.1.7 _Subject Area_

Analysis of variance shows that the subject area in which a searcher specializes has no significant effect on the average number of search keys the searcher selects (F(3, 43) = 1.09, NS).

4.2 Search-Keys Ratio

The search-keys ratio (the percentage of free-text terms selected) measures the degree to which free-text terms were used in a search, and the general preference of a searcher in the selection of search keys. Because the second objective of this study was to find the factors that affect the selection of search keys, this variable is central to the study.

The search-keys ratio associates with: the number of databases, the moves ratio, the subject area, and the environment for science searchers.

4.2.1 The Number of Databases

The variables "search-keys ratio" and "the number of databases" are directly related (search level: $r(279) = .277$, $p < .01$; person level: $r(45) = .414$, $p < .01$). That is:

> **Searches which require several databases, and searchers who habitually search several databases for a request, are likely to use more free-text terms than descriptors.**

This correlation is expected: A search that spans a number of databases is likely to include more free-text terms than descriptors because it is time consuming to look for descriptors for each database. Similarly, searchers who usually search a number of databases for each request are likely to develop a habit of using more free-text terms than descriptors for the same reason.

This association, however, warrants an examination of the causal relationships. While searchers are free to choose whether to enter free-text terms or descriptors, the number of databases to search for a request is determined by the distribution of information among the databases--it is a given. Undoubtedly, one may claim that searcher's preference of search keys can determine the number of databases to search because it is plausible to assume that searchers who feel comfortable searching with free-text terms would move from one database to another more easily than searchers who prefer to use descriptors. But even then, free-text searchers would change databases only when it is required for the success of a search. The causal relationships is, therefore, clear:

> **Having to search several databases for a request induces the use of free-text terms.**

4.2.2 Moves Ratio

The variable "search-keys ratio" directly relates to the variable "moves ratio" (search level: $r(279) = .184$, $p < .01$; person level: $r(45) = .434$, $p < .01$). Looking at the searching style of searchers, this correlation shows that:

> **Operationalist searchers prefer to use free-text terms and conceptualist searchers prefer to use descriptors.**

4.2.3 Subject Area

Analysis of variance shows that the variable "subject area" correlates with the "search-keys ratio" ($F(3, 43) = 13.16$, $p < .01$). On the average, searchers of medical literature used free-text terms 34.33% of the time, those of social sciences and the humanities 38.75%, searchers of general literature used 56.58%, and science searchers used free-text terms 75.78% of the time. A post-hoc test shows that the difference lies between science and both medicine and the social sciences searchers. That is:

> **Science searchers are more likely to use free-text terms than their colleagues who specialize in other subject areas.**

This finding presents itself as an evidence that supports common knowledge: It has been long assumed that searches in the scientific literature do not require the use of controlled vocabulary because the scientific terminology itself is already controlled. Note that this argument is not completely valid because it ignores the process of indexing which is performed mostly with controlled vocabulary, but which accomplishes additional functions such as assigning explicit terms to represent concepts which are only implicit in the text.

However, even if accepted, this argument would not be a valid explanation for this finding because of the difference between science and medical searchers. Medical terminology shows the same degree of control as science terminology, yet medical searchers used the smallest proportion of free-text terms while science searchers used the largest proportion. That is, while the fact remains that science searchers use more free-text terms than other searchers, the degree of control in the science terminology does not explain this phenomenon. Therefore, the degree to which a subject terminology is controlled is not the most important factor to determine the selection of search keys.

4.2.4 The Individual Databases

To further examine the effect of subject area on the selection of search keys, each database was assigned a subject category and the percentage of free-text terms used was calculated. Analysis of variance shows that for all searches the subject area significantly affects the proportion of free-text terms entered in a database ($F(3, 30) = 5.24$, $p < .01$). For databases in medicine and the biosciences, 36.06% of the search keys were free-text terms, 48.17% for databases in the social sciences and the humanities, 67.32% for multidisciplinary databases, and 77.90% of the search keys entered in the science and technology databases were free-text terms. A post-hoc test revealed significant difference between medicine and the sciences.

Although searchers sometimes approach databases that are outside their subject expertise--and generalist searchers may search databases in any subject--this finding reinforces the conclusion that the subject area significantly affects the selection of search keys.

4.2.5 Environment

The nature of the environment, across all subject areas, has no significant effect on search-keys ratio ($F(2, 44) = .69$, NS). However, analysis of variance shows that for those who search the scientific literature, the searcher's environment has a significant effect on the search-keys ratio ($F(1, 21) = 7.43$, $p < .05$). Science searchers who typically answer requests that address practical problems use free-text terms 86.84% of the time; those who typically search for theoretical requests use free-text terms 67.28% of the time. That is:

> **Science searchers who typically answer practical questions are more likely to use free-text terms than science searchers who usually address theoretical problems.**

The finding that environment in general does not affect the search-keys ratio but has an effect within science searching shows that the subject area has a larger effect on the selection of search keys than the nature of the requests searched.

However, it is plausible to speculate that within each subject area, practical questions encourage the use of free-text terms. The failure of this study to find such association for subject areas other than the sciences can be attributed to deficient sampling: the samples of searchers within these subject areas are too small and not representative enough. This result suggests, therefore, that:

> **Within a subject area, the nature of a request--whether practical or theoretical--may affect the percent of free-text terms selected.**

4.2.6 The Number of Moves

The Pearson Product-Moment Correlation test shows that the number of moves does not significantly relate to the search-keys ratio (search level: $r(279) = .104$, NS; person level: $r(45) = -.030$, NS). Similarly, the variables "precision moves," "recall moves," and "recall tendency" do not significantly relate to the search-keys ratio ($r(279) = -.023$, NS; $r(279) = .055$ NS; search level: $r(279) = .060$, NS, and person level: $r(45) = .104$, NS, respectively). Therefore:

> **Interaction during a search does not increase the proportion of free-text terms. Similarly, interactive searchers use the same proportion of free-text terms as searchers who are less interactive.**

Coupled with the finding that the search-keys ratio is not associated with the number of search keys (Section 4.1.4), this result is somewhat unsettling. It is sound to assume that the mechanics of the search process itself would determine the ratio of free-text terms. However, our results show that neither the number of moves per search nor the number of search keys correlates with the search-keys ratio.

While it seems plausible to conclude that interactive searchers use a large proportion of free-text terms, or that the increase in the number of search keys is always supported by adding free-text terms, our data do not support this conclusion.

4.3 Thesaurus Look-Ups

This variable measures the percent of terms entered without consulting a thesaurus. This is an important variable because nothing is gained when a searcher avoids consulting a thesaurus, and much could be lost. Further, this is not an obscure phenomenon: 37% of the free-text terms selected to search databases with indexing were picked without thesaurus consultation.

In addition, it is sound to assume that consulting a thesaurus is part of a searcher's searching style. This assumption is supported by various kinds of data. First, of the 47 searchers in this study, 32 avoided thesaurus consultation less than 20% of the times they entered free-text terms, and 5 searchers did so more than 80% of the time. That means that a total of 37 searchers exhibited a clear preference with regard to thesaurus consultation, and, moreover, that most prefer to consult a thesaurus. Second, among the reasons cited for this option (option [P]), 30% stemmed from general beliefs that searchers held, and 57% related to the databases that the searchers used regularly (see Table 2). That is, 87% of the reasons for not consulting a thesaurus stemmed from general practice, and were not related to the specific requests searched.

This variable is also important for the design of intermediary expert systems. One of the functions such an expert system could perform rather easily would be to encourage searchers to consult a thesaurus and to support them in this pursuit. It is important, therefore, to identify the factors that lead searchers to avoid using a thesaurus.

The data show that the frequency of entering search keys without consulting a thesaurus correlates with: the number of databases, the moves ratio, the subject area, the number of search keys, the number of moves, and the search-keys ratio.

4.3.1 The Number of Databases

Thesaurus look-ups and the number of databases required for a search are directly related to one another (search level: $r(279) = .294$, $p < .01$; person level: $r(45) = .397$, $p < .01$). This association should have been expected because a multi-database search was cited as a reason for not consulting a thesaurus over a quarter of the times when database-related reasons were mentioned for this option (reason [PD2] in Table 2). This association shows that:

> **The larger the number of databases to be searched per request, the more likely is the searcher to avoid consulting a thesaurus.**

Since searchers used the reason of multi-database search to explain their decision to avoid thesaurus consultation 13% of the times they elected this option (derived from Table 2), and since thesaurus consultation is a matter of searching style, the effect of the number of databases on thesaurus consultation deserves a special attention. While some searchers may feel comfortable using several databases for a search because they habitually refrain from consulting a thesaurus, for most:

> **Having to search several databases for a request induces entering free-text terms without consulting a thesaurus.**

4.3.2 Moves Ratio

Thesaurus look-ups relate directly to the moves ratio (search level: $r(279) = .167$, $p < .01$; person level: $r(45) = .413$, $p < .01$). That is:

> **Operationalist searchers are more likely to avoid consulting a thesaurus than conceptualist searchers.**

This conclusion agrees with the previous finding that operationalist searchers prefer to use free-text terms and conceptualist ones prefer descriptors (section 4.2.2).

4.3.3 Subject Area

The subject area of searching has a significant effect on the frequency with which a thesaurus is not consulted ($F(3, 43) = 3.51$, $p < 0.05$). The average frequencies of entering search keys without consulting a thesaurus for each subject area are revealing. No medical librarian in the study's sample ever entered a free-text term without checking a thesaurus, but searchers in the social sciences refrained from consulting a thesaurus 12.87% of the time. Next are general searchers with 29.32% and science searchers with 31.65%. Therefore:

> **Science searchers are more likely to enter free-text terms without consulting a thesaurus than searchers in other subject areas.**

Here again, the conclusion concurs with a previous finding: science searchers are more likely to use free-text terms than their colleagues (section 4.2.3).

Further, generalist searchers (who habitually search several subject areas) entered a considerably larger number of search keys without consulting a thesaurus than did their peers in the social sciences and medicine. This phenomenon can be explained by the fact that generalists search a relatively large number of distinct databases through all their searches. It is plausible to assume, then, that generalists search the largest number of distinct databases, even though this factor was not measured in this study. Having to use a large variety of databases, it is difficult for generalists to familiarize themselves with the thesauri of these databases and they are, therefore, more likely to refrain from using a thesaurus.

4.3.4 The Number of Search Keys

The variable "thesaurus look-ups" directly relates to the number of search keys only on the search level ($r(279) = .359$, $p < .01$), but not on the person level ($r(45) = -.164$, NS). That means that if searchers decide to increase the number of search keys for a particular request, they are likely to add terms without consulting a thesaurus, but searchers who habitually use a large number of search keys consult a thesaurus in the same frequency that other searchers do.

Since thesaurus consultation is a matter of searching style, and since it does not relate to personal inclination in the number of search keys used, the association between thesaurus look-ups and the number of search keys is induced by the nature of specific requests. The causal relationship is clear:

> **Requests that require a relatively large number of search keys lead searchers to enter search keys without consulting a thesaurus.**

This conclusion is also supported by the fact that among the request-related reasons for this option, the most common reason cited (61% of the request-related reasons, see Table 2) was that the searcher had no time to consult a thesaurus because the terms were added while online--a direct relationship between thesaurus look-ups and the number of search keys.

Another reason for not consulting a thesaurus was its unavailability to searchers (23.43% of database-related reasons for this option, see Table 2). Accounting for the instances in which this reason was used, this association generates an additional observation:

> **Thesaurus unavailability may increase the number of search keys used in a search.**

At first glance, this conclusion seems to contradict the finding that searchers who prefer to use free-text terms use, on the average, the same number of search keys as those who prefer to enter descriptors. It should be pointed out, though, that the association between thesaurus unavailability and the number of search keys is significant only on the search level. That is, when searchers who prefer free-text terms consult a thesaurus for a request, they are not likely to enter more search keys than their peers who prefer descriptors. Both types of searchers, however, are likely to increase the number of search keys when they search databases without having the pertinent thesaurus available. This increase can be explained by the fact that when a thesaurus is not available searchers are likely to examine terms that occur in the text of retrieved citations--whether descriptors or free-text terms. Such explorations usually result in entering additional search keys.

4.3.5 The Number of Moves

The number of moves relates directly to thesaurus look-ups only on the search level ($r(279) = .318$, $p < .01$), but not on the person level ($r(45) = .003$, NS). That is, when an individual request requires a relatively large number of moves, searchers are likely to avoid consulting a thesaurus, but interactive searchers do not necessarily avoid using a thesaurus more frequently than the average. Since the association between thesaurus look-ups and the number of moves is induced by the specific requests, it is clear that:

> **Interactive searches cause searchers to avoid consulting a thesaurus.**

Further, if we consider thesaurus unavailability, this association suggests that:

> **Thesaurus unavailability increases the need for interaction during the search.**

4.3.6 Search-Keys Ratio

The variables "thesaurus look-ups" and "search-keys ratio" directly relate to one another (search level: $r(279) = .299$, $p < .01$; person level: $r(45) = .660$, $p < .01$). This association is trivial, however, because it is obvious that searchers who prefer to use descriptors are more likely to consult a thesaurus (this is where they get their descriptors) than searchers who prefer to enter free-text terms.

4.3.7 Unrelated Variables

A number of variables do not correlate with "thesaurus look-ups." This lack of association is not particularly revealing, and, therefore, it merely reported here but not discussed.

The environment of searching does not significantly affect thesaurus look-ups ($F(2, 44) = 2.15$, NS). Similarly, thesaurus look-ups are not related to: "precision moves" ($r(279) = -.010$, NS); "recall moves" ($r(279) = .019$, NS); or to "recall tendency" (search level: $r(279) = .048$, NS; person level: $r(45) = .202$, NS).

To summarize, the frequency of entering free-text terms without consulting a thesaurus is: affected by the subject area, directly related to the number of databases searched and to the percentage of operational moves. In addition, thesaurus look-ups are related to the number of search keys and to the search-keys ratio only on the search level.

4.4 The Number of Databases

This variable represents the frequency of changing databases, as measured by the number of times the move Add 5 (move to a new database) was made. Like the variable "thesaurus look-ups," the number of databases is of special importance because using several databases for a single request is often necessary for the success of a search, whether or not the searcher feels comfortable changing databases. It is useful, therefore, to examine the effect of multi-database searching on the selection of search keys and on other aspects of searching behavior.

The number of databases is associated with the subject area, the number of moves, and with the moves ratio.

4.4.1 Subject Area

The subject specialty of a searcher has a significant effect on the average number of databases the searcher uses per search ($F(3, 43) = 2.25$, $p < .1$). On the average: medical librarians add .33 databases to their searches; searchers in the social sciences add 1.11 databases; generalists add 1.48; and science searchers add 1.64 databases per search. Therefore:

Science searchers are more likely to use several databases per search than searchers in other subject areas.

4.4.2 Moves

The number of moves directly relates to the number of databases (search level: $r(279) = .592$, $p < .01$; person level: $r(45) = .631$, $p < 0.01$). At first sight, this association seems trivial since changing a database by itself is a move. To create a more meaningful association, the number of moves was redefined to exclude the move of changing databases (Add 5). The total number of moves without Add 5 directly relates to the number of databases used in a search ($r(279) = .309$, $p < 0.01$). Similarly, the average number of moves, excluding Add 5, made by a searcher per search directly relates to the average number of databases used per search ($r(45) = .406$, $p < .01$). That is:

Interactive searches are more likely to require use of several databases than less interactive searches. Similarly, interactive searchers are more likely to use several databases per search than their peers who are less interactive.

In addition, the number of databases directly relates to the number of precision moves ($r(279) = .168$, $p < .01$), to the number of recall moves ($r(279) = .596$, $p < .01$), and to the percent of recall moves on the search level ($r(279) = .185$, $p < .01$). On the person level, the number of databases does not relate to "recall tendency" ($r(45) = .203$, NS). That is, searchers who are usually concerned with recall do not habitually use more databases than searchers who are usually concerned with precision.

4.4.3 Moves Ratio

The percentage of operational moves directly relates to the number of databases searched (search level: r(279) = .312, p < .01; person level: r(45) = .370, p < .01). This correlation is trivial, however, because changing databases by itself is an operational move. To examine whether operationalist searchers are more likely than their conceptualist counterparts to use several databases, the moves ratio was redefined as the number of operational moves--not counting the move Add 5--divided by the total number of moves without Add 5. The new moves ratio does not significantly correlate with the number of databases (search level: r(279) = -.058, NS; person level: r(45) = .032, NS). This test reaffirms the conclusion that this association is of no significance.

4.4.4 Environment

The search environment has no significant effect on the number of databases searched (F(2, 44) = .08, NS). That is, the nature of a request, whether practical or theoretical, is likely to have no effect on the number of databases searched.

4.5 The Number of Moves

The number of moves does not correlate with any of the remaining variables. While it is directly related to "precision moves" ($r(279) = 0.483$, $p < .01$), to "recall moves" ($r(279) = .905$, $p < .01$), and to "recall tendency" on the search level ($r(279) = .254$, $p < .01$), these associations are trivial, as explained before. The number of moves does not, however, relate to "recall tendency" on the person level ($r(45) = 0.172$, NS). That is, interactive searchers are concerned with recall to the same degree as are other, less interactive searchers.

In addition, the percentage of operational moves does not relate to the number of moves (search level: $r(279) = -.022$; NS, person level: $r(45) = -.004$, NS). That is:

> **Operationalist and conceptualist searchers are interactive to the same degree.**

Further, the number of moves is affected neither by the subject area ($F(3, 43) = .15$, NS), nor by the environment ($F(2, 44) = 1.95$, NS). That is:

> **The subject area in which a searcher specializes or the environment in which the searcher works do not affect the searcher's level of interaction.**

The effect of the environment was measured within each subject area. A significant effect was found for medical searchers ($F(2, 5) = 99.67$, $p < .01$). Medical librarians who answer mostly practical questions made an average of 3.97 moves per search, while those who usually answer theoretical questions made an average of 18.78 moves per search. This drastic difference is probably the result of the extreme difference between the practical and theoretical settings for the medical searchers who participated in the study. While hospital librarians composed the largest part of this sample, regulatory agencies were the only theoretical environment for this study. Searches in these agencies require the highest degree of recall, and therefore, may require interaction on a level that is much higher than the average. It is premature, therefore, to conclude that the environment affects the level of interaction.

4.6 The Moves Ratio

The moves ratio reflects the searching style of a searcher whether operationalist or conceptualist. It is defined here as the percentage of operational moves made on the average by a searcher and in an individual search.

The moves ratio is affected by the subject area, and it is associated with the variables "precision moves" and "recall tendency."

4.6.1 Subject Area

The subject area in which a searcher specializes has a significant effect on the "moves ratio" ($F(3,43) = 6.31$, $p < .01$). Medical searchers selected 45.11% of their moves to be operational, searchers in the social sciences made 50.54% operational moves, science searchers made such moves 76.03% of the time, and general searchers selected 79.28% of their moves to be operational. A post-hoc test found that the difference lies between general searchers and both medical and social-sciences searchers, as well as between medical and social-sciences searchers. That is:

> **Science searchers and searchers who have no subject specialty are more likely to make operational moves than their colleagues in other subject areas.**

The large percent of operational moves among generalist searchers can be explained by the nature of their task: They are called upon to answer requests in a large variety of subjects. Unlike searchers who specialize in one subject area, their knowledge of the subject of a request is usually limited. This limitation prevents them from making conceptual moves because conceptual moves by their nature require some subject knowledge: they are moves that change the meaning of a request. A person who is familiar with the subject of a request is more likely to feel comfortable modifying its meaning for the purpose of a search than a person who has little expertise in the subject matter.

While the tendency to make operational moves among generalist searchers is inherent to the nature of their searching, finding this tendency among science searchers is new data about searching behavior.

4.6.2 Moves

The "moves ratio" is directly related to the number of precision moves in a search ($r(279) = .240$, $p < .01$), but is not significantly related to "recall moves" ($r(279) = .050$, $p < .01$). That is:

> **Precision moves are more likely to be operational than conceptual ones.**

Recall tendency, on the other hand, is directly related to the "moves ratio" on the search level ($r(279) = .141$, $p < .05$), but not on the person level ($r(45) = -.186$, NS). This means that while searchers who are usually concerned with recall do not have a particular style of searching, a request that requires more recall moves than any other moves is likely to be searched with operational moves. As discussed earlier, this conclusion might be a result of the large frequency with which the move Add 5 was made: it is both an operational move and a move to improve recall. Therefore, the association between "recall tendency" and the "moves ratio" is not significant to the study of online searching behavior because it might have been induced by the need to search several databases, and because it is manifested only on the search level.

4.6.3 Environment

The environment has no significant effect on the moves ratio (F(2, 44) = 1.24, NS). That is:

> **The environment in which a searcher works has no effect on the searching style of the searcher.**

This conclusion was found to hold for the environments within each subject area.

4.7 Recall Tendency

Recall tendency (the percentage of recall moves made by a searcher across all searches) reflects the degree to which a searcher is usually concerned with recall.

Analysis of variance found that recall tendency is significantly affected neither by the subject area ($F(3, 43) = .52$, NS), nor by the environment across subject areas ($F(2, 44) = 2.83$, NS). The same analysis for environments within each subject, however, revealed that the environment within the sciences significantly affects recall tendency ($F(1, 21) = 7.29$, $p < .05$). Science searchers in theoretical environments made recall moves 74.51% of the time, while those in practical environments made such moves 54.94% of the time. That is:

> **Science searchers who work in theoretical environments are more likely to be concerned with recall than their colleagues in practical environments.**

This searching behavior of science searchers can be looked upon as a reflection of the nature of science requests. Therefore, it is plausible to assume that:

> **A scientific request that is theoretical in nature may require more recall moves than a request that is practical.**

A summary of the findings reported in this section is presented in Table 6.

Table 6. Summary of the factors affecting searching behavior

	1	2	3	4	5	6	7	8	9	10	11
1. The no. of search keys	*	NS	.01	.01	.01	NS	.01	.01	NS	NS	.01
2. Search-keys ratio	NS	*	.01	.01	NS	.01	NS	NS	NS	.01	.05
3. Thesaurus look-ups	.01	.01	*	.01	.01	.01	NS	NS	NS	.05	NS
4. The no. of databases	.01	.01	.01	*	.01	.01	.01	.01	.01	.1	NS
5. The no. of moves	.01	NS	.01	.01	*	NS	*	*	*	NS	NS
6. Moves ratio	NS	.01	.01	.01	NS	*	.01	NS	.01	.01	NS
7. Precision moves	.01	NS	NS	.01	*	.01	*	*	*	*	*
8. Recall moves	.01	NS	NS	.01	*	NS	*	*	*	*	*
9. Recall tendency	NS	NS	NS	.01	*	.01	*	*	*	NS	0.5
10. Subject area	NS	.01	.05	.1	NS	.01	*	*	NS	*	*
11. Environment	.01	.05	NS	NS	NS	NS	*	*	.05	*	*

5. SUMMARY AND CONCLUSIONS

The purpose of this study was to uncover the rules used by online searchers for the selection of search keys, whether free-text terms or descriptors, and to represent these rules in a formal model that could be used in the construction of intermediary expert systems.

The case study method with controlled comparison was used to analyze the data collected through observation of online searchers performing their regular, job-related searches. The study's participants were experienced searchers who were selected from a wide spectrum of subject specialties and from various settings.

Data analysis was based on two existing models. The first was the original Selection Routine, a decision tree that presented the rules used to select search keys by eight searchers in a previous study. The second was a list of moves--modifications in search strategies--based on observing the searching behavior of the same eight searchers. The list is divided into two types of moves: operational moves which keep the meaning of a request unchanged; and conceptual moves which change the meaning of a request. Within each type, the moves are presented in three groups: precision moves; recall moves; and moves to increase both precision and recall.

This study included 39 searchers whose searching behavior was analyzed in order to expand both models. These searchers were also asked to explain their reasons for each selection of a search key.

Data analysis involved measuring the frequency with which: (1) each type of search key was selected; (2) each move was selected; and (3) a reason was cited to explain the selection of a search key. Further, the statistical associations among eleven variables were examined. These variables are: (1) the number of search keys selected for a search; (2) search-keys ratio (the percentage of free-text terms compared to the total number of search keys); (3) thesaurus look-ups (the frequency with which a thesaurus was not consulted); (4) the number of databases used per search; (5) the number of moves made in a search; (6) moves ratio (the percentage of operational moves compared to the total number of moves); (7) the number of precision moves made in a search; (8) the number of recall moves made in a search; (9) recall tendency (the percentage of recall moves compared to the total number of moves); (10) the subject area in which a searcher specializes; and (11) the environment in which a searcher works.

The statistical analyses included data from the 39 study searchers as well as from the eight original searchers (a total of 47), and were performed on two levels: the search level, in which each search was considered as an instance (281 instances); and the person level, in which data from all the searches performed by one person were aggregated to represent a single instance (47 instances).

The study had three objectives: (1) to refine and validate the Selection Routine; (2) to explore the effect of searching behavior on search-key selection; and (3) to test the applicability of the case study method to the extraction of knowledge from multiple experts.

This chapter summarizes the results of the study that are relevant to each of the objectives, and examines the applicability of the findings to the design of both bibliographic databases and intermediary expert systems.

5.1 The Selection Routine

The analysis of searching behavior performed for this study modified the Selection Routine. It discovered a new condition (the searcher did not consult a thesaurus), and added a few options to existing conditions.

The modified Selection Routine can now be used to develop the set of rules for a rule-based intermediary expert system, even though the set is not complete: A number of options have no reasons to explain their selection. For example, when a common term is mapped to a descriptor, the Selection Routine provides two options: use descriptors [A], or use free-text terms [B]. While option [B] was selected only when the term was used as a limiting factor, no reasons were given by searchers for the instances in which they decided to enter a *descriptor* as a limiting factor. Thus, the reasons for preferring one or the other of these options are not clear.

Similarly, no reasons were elicited for the options [D], [E], [M], [N], [Z10], and [Z11]. Thus, while these options can be recommended for use by an intermediary expert system, such a recommendation cannot be based on attributes of the database searched, the request, or on attributes of the user.

Note, however, that the searchers who participated in the study selected these options very infrequently--several of these options were selected only once. This phenomenon may indicate that these options are not as viable as the other options. Therefore, they may be suggested for use by an intermediary expert system only as the last resort.

An alternate conclusion is that these options are indeed viable but that the conditions for which they are useful occur in low frequency. Therefore, to cover a larger range of possible conditions, future studies should address these options, the reasons for their selection, and their applicability to the design of intermediary expert systems.

In addition, the frequency and the reasons for selecting each option were examined. This analysis revealed several important patterns in online searching behavior.

5.1.1 The Selection of Options

Searchers selected the most straighforward options (that is, to enter a descriptor when a term is mapped to a descriptor exactly, and to enter a free-text term without consulting a thesaurus or when a term cannot be mapped to a descriptor) about 70% of the time (section 3.2.2).

This phenomenon can be accounted for by two possible explanations. First, 70% of the terms selected for the requests submitted to the study's searchers did not present any terminological problems and therefore could be entered directly. Second, searchers' tendency was to avoid selecting options that are not straighforward.

Note, however, that 20% of the times searchers selected search keys, they entered free-text terms because they did not consult a thesaurus (option [P] in table 2), and this at times because it was unavailable (19% of the database-related reasons, section 3.2.3). Further, the decision of whether or not a term presents terminological problems is often subjective: it is determined by the searcher's perception. Therefore, while both explanations account for this phenomenon, the searchers perceptions of terminological complexity and the availability of thesauri seem to be prominent motivations to select straightforward options.

This finding is relevant to the design of intermediary expert systems. One may claim that if such systems are to accurately duplicate searching behavior of human intermediaries, extremely simple systems--based on a single rule--could still "succeed" 70% of the time in selecting the "right" type of search key. On the other hand, this finding demonstrates the potential power of intermediary expert systems to enhance searching. These systems could routinely look up the pertinent thesaurus for each term and thus eliminate the condition in which a searcher does not know if a term is mapped to a descriptor. Further, they could provide tools that would simplify the resolution of terminological problems. Thus, with the help of an intermediary expert system, the frequency in which straightforward options are selected because of searching difficulties could be reduced.

5.1.2 Thesauri Quality and Availability

The important role of databases in the selection of search keys is strongly demonstrated by the findings of this study. Most notable is the finding that searchers consulted a thesaurus for 80% of the search keys they selected, and given a choice, they selected descriptors about 50% of the time (section 3.2.1). Further,

--searcher's selection of search keys was most often determined by the database they were searching (48% of the time, section 3.2.3),

--when searchers explained their selection with a database-related reason, 25% of the time they observed that a term would not be in the thesaurus, 19% of the time they explained that a thesaurus was unavailable to them, and 18% of the time they claimed that they do not trust the descriptors and/or the indexing (section 3.2.3),

--distrust of descriptors and/or indexing explained 16% of the instances in which searchers entered terms without consulting a thesaurus (section 3.2.2),

--the study's results suggest that searchers do not exercise terminological control in searching (section 4.1.4), and

--the study's results support the hypothesis that searchers are more likely to enter terms without consulting a thesaurus when they search databases for which they use descriptors infrequently than when they search databases for which they usually use descriptors (section 3.2.4).

Therefore, the quality and availability of thesauri are critical factors in the selection of search keys. Further, the results of this study show that better quality in thesauri and indexing--and greater availability of these tools--are badly needed.

5.1.3 The Concern with Recall

Recall, which measures the completeness of the information retrieved, is of special concern in information science research. This concern is based on the findings of experiments in online retrieval: Most experiments have resulted in relatively low recall scores. For example, in a study completed recently by Saracevic and Kantor, precision was 57% for all searches but recall was only 22% [Saracevic & Kantor, 1988]. As the authors explain, these ratios agree with results of other studies.

The searchers who participated in the study attempted most often to increase recall:

--among the options that are not straightforward, over a half were selected to enhance recall (section 3.2.2),

--among the request-related reasons for the selection of search keys, the need to enhance recall was the most frequent reason (35% of request-related reasons, section 3.2.3), and

--the number of moves to increase recall was almost double the number of moves to increase precision (section 3.3.1).

The low recall scores obtained in experiments has often raised the concern that searchers in general do not consider recall to be an important factor, or that they prefer to avoid the extra effort that is presumably required to increase recall. On the contrary, the findings of this study show that searchers do consider recall an important factor when they select search keys, and when they modify search strategies with moves.

The discrepancy between the findings of this study and the low recall scores obtained in online-searching experiments can be partially attributed to the study methods used. While searchers who participate in experiments search requests under artificial conditions, this study examined searchers answering real-life requests submitted by users to whom the searchers are accountable. It is possible, therefore, that the searching observed in this study was guided by a level of recall-consciousness that is higher than the one exhibited in experiments carried out to study online searching behavior. While this an important observation for future online-searching experiments, it is difficult to substantiate this conclusion because recall ratios were not measured in this study, and thus no comparison between recall scores obtained in this study and those measured in experiments can be made.

On the other hand, this discrepancy could be explained by the observation that current bibliographic databases do not provide for high recall ratios. In other words, regardless of searcher's experience or searching style, it is difficult to achieve recall scores that are satisfactory when using the current bibliographic databases.

This conclusion highlights the importance of recall in online retrieval. Designers of both databases and intermediary expert systems should pay special attention to means to improve recall and provide tools that support searchers' attempts to enhance recall. Moreover, this conclusion calls for further explorations to discover new ways to improve the recall of retrieved sets.

5.2 Factors Affecting Searching Behavior

The second objective of the study was to determine the factors that affect the selection of search keys and other aspects of searching behavior. This objective was addressed on four levels. First, the characteristics of the searchers who are likely to prefer free-text terms was determined. Second, the effects of factors that are typical of the searching behavior of a searcher were examined. Third, characteristics of requests that may affect searching behavior were scrutinized. Fourth, the effect on searching behavior of decisions usually made by designers of databases and intermediary expert systems was analyzed.

5.2.1 The "Free-Text" Searcher.

The variable "search-keys ratio," when analyzed on the person level, measures the degree to which a searcher prefers to use free-text terms. Results reported in section 4.2 show that a profile of the searchers who use free-text terms more often than other searchers can now be constructed. These searchers are likely to have these characteristics:

--be operationalist searchers,

--be science searchers,

--if, as science searchers, they usually answer practical requests, they will use still more free-text terms,

--need to search several databases for each request, and

--will have developed a habit of entering terms without consulting a thesaurus.

Note that searchers who prefer to enter free-text terms _do not_ enter more search keys than those who prefer descriptors (section 4.1.4), nor are they more interactive than their counterparts (section 4.2.6).

The nature of the "free-text" searcher as described here raises the question: Is the preference of free-text terms an inherent attribute? That is, is it determined by factors such as cognitive style or personality traits? Answering this question is significant to research in online searching behavior because most of this research has focused on inherent characteristics of searchers.

The results of this study show that inherent attributes have some effect on habitual preference in the selection of search keys: operationalist searchers prefer to use free-text terms. These results show, at the same time, that the tendency to prefer free-text terms is encouraged by the realities of searching: by the subject area, the environment, the number of databases, and by the availability and quality of thesauri. This conclusion is supported by another finding: only 20% of the reasons for selecting a search key stemmed from habitual searching behavior (section 3.2.3). That is, the selection of search keys is most frequently determined by the specific requirements and constraints of a search, and the effect of inherent searching behavior on this selection is less extensive.

5.2.2 Factors Typical of Searching Behavior

The variable "search-keys ratio" measured the degree to which searchers prefer to use free-text terms. Two more aspects of online searching behavior are embodied in the variables tested in this study. First, the typical level of effort a searcher put into the completion of a search. Second, the searching style of a searcher, whether operationalist or conceptualist.

The first aspect--the average effort--can be measured by the number of search keys entered, by the number of moves made per search, and by the number of databases used per search. The first two variables are associated with one another but are not associated with any other aspect of online searching behavior. That is, regardless of whether they prefer free-text terms or descriptors, and regardless of whether they are operationalist or conceptualist searchers, some searchers routinely put more effort to their searches than others: searchers who tend to be interactive during an online session are likely to use more terms than their peers who are less interactive (section 4.1.1). In addition, interactive searchers are likely to use more databases than their less-interactive colleagues (section 4.4.2).

The searching style of a searcher, whether operationalist or conceptualist, may also affect the selection of search keys and other aspects of searching behavior. In this study, the variable "moves ratio" measured the degree to which a searcher was operationalist, as determined by the moves the searcher made. The results show that operationalist searchers:

--use free-text terms more frequently (section 4.2.2),

--are more likely to avoid consulting a thesaurus (section 4.3.2),

--are more likely to answer science or general questions (section 4.6.1), and

--are more likely to make precision moves than conceptualist searchers (section 4.6.2).

That is: Although only 25 of the 47 searchers exhibited a strong commitment to one type of moves (operational or conceptual, section 3.3), operationalist searchers differ from their conceptualist peers in their preference for the type of search keys, their habits relating to thesaurus look-ups, in their subject speciality, and in their concern about precision.

5.2.3. The Effect of Requests on Searching Behavior

The nature of a request is central to the search process. Ideally, the search process should be determined by the nature and requirements of the specific request. It is significant, therefore, to examine the actual effect that requests had on the searching behavior of the study's participants, and in particular, on their selection of search keys. This examination is guided by some general conclusions, as well as by specific variables, which address request characteristics.

One measure of the degree to which individual requests affect the selection of search keys is the percentage of request-related reasons given to explain the selection of search keys. The study's searchers referred to requirements put by requests 32% of the times they explained their search-keys selection (section 3.2.3). While one wants to hope for a higher percentage, it should be remembered that 48% of the reasons related to constraints of the databases. It is plausible to assume, therefore, that with more flexible structure, and with better availability of searching tools, searchers will give request characteristics higher priority.

On the other hand, requests introduced the largest variability to reasons for the selection of search keys (section 3.2.3). This means that in comparison to database constraints and to individual searching habits, request requirements are the least predictable.

Therefore, designers of databases should provide for higher flexibility in searching so searchers could adjust their search strategies to the requirements of individual requests. Further, intermediary expert systems should be designed to explore the nature of requests so they can make informed decisions about the selection of search keys.

Although the effect of the characteristics of a request is clearly demonstrated by this study, it is revealing to examine some factors that are free of this effect. First, contrary to common belief, high-recall requests do not require an increased number of search keys (section 4.1.6). Second, the nature of a request seldom determined whether or not a searcher would consult a thesaurus: only 12% of the reasons for not consulting a thesaurus were related to requests (section 4.3.7).

In addition to these general observations, the nature of requests is reflected in several variables. These are: the number of search keys; the number of moves; subject area; and environment. The effect of requests on searching behavior can be determined, therefore, by statistical associations when measured for these variables on the search level--where each search was considered an individual instance.

One should be cautious, though: instances of individual searches are not independent because every five searches were performed by the same person. That is, effects that are detected might be induced by the searchers, rather than by the requests. Therefore, most of the relationships established in this study do not constitute evidence that requests affect searching behavior, but they suggest possible association.

The first variable--the number of search keys--is used here to reflect the terminological difficulty of a request. While the number of search keys is primarily determine by the number of components a request has, the latter number is usually limited: almost no request includes more than four components. Therefore, having a relatively large number of search keys in a query formulation is usually the result of the need to represent each component with several search keys--a situation that is caused by terminological difficulties.

Requests with terminological difficulties generate distinct searching behavior. First, they result in increased interaction, since the number of search keys is associated with the number of moves (section 4.1.1). This association, however, holds also on the person level and it is, therefore, suggestive only.

Second, requests with terminological difficulties lead searchers to enter search keys without consulting a thesaurus (section 4.3.4), as do requests that require a large number of moves (section 4.3.5). These associations are particularly significant because they do not hold on the person level. That is, searchers who habitually enter a large number of search keys, or those who are typically interactive, do not necessarily avoid consulting a thesaurus more frequently than their peers.

Thus, requests with terminological difficulties induce a certain pattern of searching: they require more interaction than other requests, an interaction during which searchers add search keys without consulting a thesaurus. The implication of this pattern to the design of databases and intermediary expert systems is clear: easy and inexpensive access to thesauri during online sessions will enhance the search process.

The other two variables that describe characteristics of requests are the "subject area" and the "environment" of searching. These variables were measured on the person level only, and therefore, they depict whether searching in a certain subject area and in a distinct environment affect searching behavior. As explained earlier, conclusions about the nature of individual requests are only suggestive.

The results of the study show that:

--science requests are searched with free-text terms more frequently than requests in other subject areas (section 4.2.3), and

--science and general requests are more likely to be searched by operational moves than by conceptual ones (section 4.6.1).

These results are not conclusive because they refer to factors that are typical of searching behavior: the preference to use free-text terms, and the searching style. To date, it is not clear whether searching the science literature causes a searcher to develop inherent characteristics, or whether searchers select the subject area in which they search according to their searching habits. Nevertheless, these results demonstrate that the subject area plays a role in searching behavior.

The variable "environment" was actually defined to characterize the nature of requests: whether practical or theoretical. These categories were assigned intuitively, and were determined by the mission of the organization in which a searcher worked. Thus, this variable is not rigorously defined.

Results of the study, on the other hand, clearly demonstrate that the nature of requests affect searching behavior: practical requests are searched with free-text terms more frequently than theoretical ones (section 4.2.5), and theoretical requests require higher recall than practical ones (section 4.7).

These findings agree with prevalent ideas. It is commonly assumed by searchers that when faced with a practical problem, users need a few articles that are highly relevant and there is no evidence to contradict this assumption. One way to achieve a high level of relevance is to require that the terms employed by the user appear in titles of articles--a practice that was frequently mentioned by searchers (56% of the searcher-related reasons, section 3.2.3). This approach is sound because terms used in practical requests are usually defined better than those in theoretical requests and present less terminological problems, and because recall is not an important factor since users are not concerned with the information they might have missed but rather with their ability to solve their problem. This approach, however, is carried out by using free-text terms--a practice that is reflected in the association between the variables "environment" and "search-keys ratio."

While still unsubstantiated hypotheses, these findings are pertinent to the design of intermediary expert systems. An expert system which mediates between end users and bibliographic databases should help users to determine whether their request is practical or theoretical. Once the nature of a request is determined, the system can make decisions about the type of search keys to be selected and about the moves that would enhance retrieval.

5.2.4 The Effects of Design Factors

Among all variables examined in this study, only one relates to design factors: the number of databases used in a search. While it may seem that searchers are free to choose the number of databases they search, their decision is determined by the distribution of information among databases rather than by their "desire" to try new databases. That is, the distribution of information among databases within a subject area determined the number of databases that were used per search. In other words, the number of databases that need to be used is a given with searchers.

Because databases are carved to fit a subject area, this observation is substantiated by the finding that the subject area affects the number of databases used per search, and in particular by the difference between medical and science searchers: the former searched an average of 1.33 databases per search and the latter an average of 2.64 per search (section 4.4.1).

Further, the results show that searching behavior is affected primarily by the databases (section 3.2.3), and that this variable correlates with the largest number of variables (Table 6). It is crucial, therefore, to examine the effect of multi-database searching on the selection of search keys.

Statistical tests revealed that having to search several databases for a request induces the use of free-text terms (section 4.2.1), and entering free-text terms without consulting a thesaurus (section 4.3.1). Further, having to search a number of databases was cited as a reason for entering free-text terms and for not consulting a thesaurus (section 3.2.3).

These findings provide evidence for the conclusion that the use of several databases causes searchers to enter free-text terms and to avoid consulting a thesaurus. This effect is obviously an impediment to searching because it limits the choices in the selection of search keys that searchers can have.

This conclusion has direct implications for the design of databases and intermediary expert systems. It is plausible to assume that if databases were more "similar" to one another, moving from one database to another would not affect the selection of search keys. More research is needed to discover which features of databases should be kept similar, and what kind of variability is desired. It is clear, however, that the findings of such research could not be implemented without standardization and cooperation in the design and production of databases.

Another approach to minimize the affect of multi-database searching on the selection of search keys is to introduce a switching language that "translates" the vocabulary of one thesaurus into another, and the vocabulary of a user into the vocabulary of a designated thesaurus. Indeed, the use of such a language has already proven to be useful [Chamis, 1988]. Such languages are designed for intermediary expert systems, so that descriptors and free-text terms can be selected by a system for each request and for every database that is to be searched without user assistance.

The conclusion that multi-database searches have an effect on the selection of search keys only emphasizes the importance of this component in intermediary expert systems: these systems should mask the differences between databases. One should remember, however, that the differences between databases are not a necessity; they are introduced most often because of commercial considerations that may or may not satisfy searching needs. It is more useful to avoid unnecessary inconsistency in database design, and to mask the necessary variability.

Thus, research should be carried out to discover which features of databases and their thesauri can be standardized without affecting their retrieval quality. The role of intermediary expert systems will then be to bridge across the necessary differences, employing switching languages and other terminological and semantic networks.

5.3 The Case Study Method

The applicability of the case study method to the extraction of knowledge from multiple experts is proven by the successful generation of formal models that describe the selection of search keys and the moves that searchers make. The use of this method in this study led to two conclusions: (1) the method of controlled comparison is useful to resolving conflicting evidence; and (2) observation and analysis of a relatively small number of searchers is sufficient to create a model that describes their searching behavior in formal terms.

The method of controlled comparison is used to explain observations that are seemingly contradictory. For example, according to the Selection Routine, searchers have two options when a single-meaning term is mapped to a descriptor through partial match: they can enter the descriptor, or they can use free-text terms for an inclusive search. These two options are not similar to one another. The reasons provided by searchers to explain their choice, however, uncovered additional factors that are used: Concern for recall may encourage searchers to use a free-text term in an inclusive mode, if possible, or it may direct searchers to enter the descriptor if it is only added to the formulation, or if it was spotted in the indexing of a relevant citation. Thus, request requirements and indexing were discovered to be factors that affect the selection of search keys.

The original Selection Routine was based on the observation of the searching behavior of eight searchers. The observation of the study's 39 searchers did not result in major modifications. Only two moves were added to the list of moves and both were used infrequently. On the other hand, a new condition was added to the original Selection Routine: a searcher does not know if a term is mapped to a descriptor. This condition was not spotted in the observations for the original Routine because that study was limited to medical librarians who never searched a request without consulting a thesaurus. This new condition was uncovered through the first searcher who was selected from another subject area.

The experience derived from using the case study method shows that limiting the sample of searchers to be observed by factors such as subject area or environment prevents the creation of a general model of searching behavior. On the other hand, only two options were added to the original Selection Routine. Thus, if one takes into account the variety that exists among searchers, the observation of a relatively small number of searchers is sufficient for the creation of formal models that describe their searching behavior.

5.4 Implications for Future Research

The findings of this study raised new questions, and point to new issues for research. Among these issues, four are relevant to research in online searching behavior in general.

First, this study demonstrates that searching behavior follows certain patterns, and that general laws govern this behavior. This conclusion is timely because most studies in online searching behavior have concluded that such laws cannot be discovered. Moreover, results of several experiments have led investigators to believe that individual variability among searchers is large enough to obscure the patterns that may exist. The results of this study, however, prove that individual variability, as expressed by general beliefs held by searchers, has the smallest effect on the selection of search keys, and possibly on other aspects of searching behavior.

It is now time to reassess the methods and techniques used in the study of online searching behavior, as well as the issues that are selected for investigation.

Second, the results of this study add to existing evidence that current systems cannot provide satisfactory recall. Although methods to improve recall are known and are in use when needed (e.g., the moves to increase the size of the set), we still do not know why recall scores on the average are much lower than precision scores. This issue should be addressed by researchers in order to discover impediments to recall, and to create design modifications that could enhance recall of retrieved sets.

Third, the requests presented by users were found to introduce the largest variability in the decisions about selection of search keys. This means that requests are the least predictable among the factors that affect searching behavior (i.e., databases and searcher's beliefs). The importance of the nature of the request to the search process has long been recognized. However, despite various attempts to discover the effect of the nature of requests on searching behavior and on the quality of retrieved sets (e.g., [Saracevic & Kantor, 1988]), no definite conclusions exist as yet. Research about the effect of requests on searching behavior should focus on request characteristics that are significant to the search process--the first step in this direction is to uncover these characteristics. For this purpose, a better understanding of the search process itself is needed.

Fourth, a large array of findings provides the evidence for the conclusion that existing databases leave much to be desired, and that some are even an impediment to useful searching practices. Further explorations are needed to discover what difficulties are encountered with databases--their thesauri and indexing--and what frustrations are experienced by searchers using them. Due to the controversy about the most efficient and cost-effective methods for information retrieval, it seems that there is no agreed-upon alternative that is superior to the current databases. Identifying flaws in existing databases, thesauri, and indexing from the searchers' point of view would provide guidelines for the design of better systems.

6. REFERENCES

Baker, Christine; Eason, Kenneth D. 1981. An Observational Study of Man-Computer Interaction Using an Online Bibliographic Information Retrieval System. *Online Review*, 5(2):121-132; April 1981.

Belkin, Nicholas J. ; Vickery, Alina. 1985. *Interaction in Information Systems: A Review of Research from Document Retrieval to Knowledge-Based Systems*. London: The British Library, 1985. (Library and Information Research Report 35)

Brooks, Helen M. 1987. Expert Systems and Intelligent Information Retrieval. *Information Processing & Management*, 23(4):367-382.

Carrow, Deborah; Nugent, Joan. 1981. Comparison of Free-Text and Index Search Abilities in an Operating Information System. *Information Management in the 1980s: Proceedings of the American Soceity for Information Science 40th Annual Meeting, September 26-October 1, 1977*. Volume 14. White Plains, NY: Knowledge Industry Publications, 1981. Pp. 131-138.

Chamis, Alice Y. 1988. Selection of Online Databases Using Switching Vocabularies. *Journal of the American Society for Information Science*, 39(3):217-218; May 1988.

Cleverdon, Cyril W. 1962. *Report on the Testing and Analysis of an Investigation Into the Comparative Efficiency of Indexing Systems*. Cranfield, England: College of Aeronautics, ASLIB Cranfield Research Project, 1962.

Cleverdon, Cyril W. 1984. Optimizing Convenient Online Access to Bibliographic Databases. *Information Services & Use*, 4(1/2):37-47.

Croft, W. Bruce. 1987. Approaches to Intelligent Information Retrieval. *Information Processing & Management*, 23(4):249-254.

Daniels, P.J. 1986. Cognitive Models in Information Retrieval--an Evaluative Review. *Journal of Documentation*, 42(4):272-304; December 1986.

Diesing, Paul. 1971. *Patterns of Discovery in the Social Sciences*. Chicago, IL: Aldine-Atherton, 1971.

Doszkocs, T.E. 1983. Automatic Vocabulary Mapping in Online Searching. *International Classification*, 10(2):78-83.

Dubois, C.P.R. 1987. Free Text vs. Controlled Vocabulary; a Reassessment. *Online Review*, 11(4):243-253; August 1987.

Ericsson, K. Anders; Simon, Herbert A. 1984. *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: The MIT Press, 1984.

Fenichel, Carol H. 1980. The Process of Searching Online Bibliographic Databases: A Review of Research. *Library Research*, 2(2):107-127.

Fidel, Raya. 1984a. The Case Study Method: A Case Study. *Library and Information Science Research*, 6(3):273-283; July-September 1984.

Fidel, Raya. 1984b. Online Searching Styles: A Case-Study-Based Model of Searching Behavior. *Journal of the American Society for Information Science*, 35(4):211-221; July 1984.

Fidel, Raya. 1985. Moves in Online Searching. *Online Review*, 9(1):61-74; February 1985.

Fidel, Raya. 1986. Towards Expert Systems for the Selection of Search Keys. *Journal of the American Society for Information Science*, 37(1):37-44; January 1986.

Fidel, Raya. 1988. What is Missing in Research About Online Searching Behavior? *The Canadian Journal of Information Science*, 12(3/4):54-61; May 1988.

Fugmann, Robert. 1982. The Complementarity of Natural Language and Indexing Languages. *International Classification*, 9(3):140-144.

Henzler, Rolf G. 1978. Free or Controlled Vocabularies: Some Statistical User-Oriented Evaluations of Biomedical Information Systems. *International Classification*, 5(1):21-26; March 1978.

Katzer, Jeffrey. 1982. A Study of the Overlap Among Document Representations. *Information Technology: Research and Development*, 1(4):261-274; October 1982.

Keen, Edward Michael. 1973. The Aberystwyth Index Languages Test. *Journal of Documentation*, 29:1-35; March 1973.

Kirby, Martha; Miller, Naomi. 1986. MEDLINE Searching on Colleague: Reasons for Failure or Success of Untrained End Users. *Medical Reference Services Quarterly*, 5(3):17-34; Fall 1986.

Krawczak, Deb; Smith, Philip J.; Shute, Steven J. 1987. EP-X: A Demonstration of Semantically-Based Search of Bibliographic Databases. In: *Proceedings of the 10th Annual International ACM SIGIR Conference in Research & Development in Information Retrieval, New Orleans, LA, June 3-5, 1987*, edited by C.T. Yu and C.J. Van Rijsbergen. New York: ACM, 1987. Pp.263-271.

Lancaster, F. Wilfrid. 1980. Trends in Subject Indexing from 1957 to 2000. In: Taylor, P.J., ed. *New Trends in Documentation and Information*. London: Aslib, 1980.

Marcus, Richard S. 1983. An Experimental Comparison of the Effectiveness of Computers and Humans as Search Intermediaries. *Journal of the American Society for Information Science*, 34(6):381-404; November 1983.

Markey, Karen; Atherton, Pauline; Newton, Claudia. 1980. An Analysis of Controlled Vocabulary and Free Text Search Statements in Online Searches. *Online Review*, 4(3):225-236; September 1980.

Miles, Matthew B.; Huberman, A. Michael. 1984. *Qualitative Data Analysis: A Sourcebook of New Methods*. Beverly Hills, CA: Sage, 1984.

Monarch, Ira; Carbonell, Jaime. 1987. CoalSORT: A Knowledge-Based Interface. *IEEE Expert*, 2(1):39-53; Spring 1987.

Oldroyd, B.K. 1984. Study of Strategies Used in Online Searching 5: Differences Between the Experienced and the Inexperienced Searcher. *Online Review*, 8(3):233-244; June 1984.

Paice, Chris. 1986. Expert Systems for Information Retrieval? *Aslib Proceedings*, 38(10):343-353. October 1986.

Parker, J.E. 1971. Preliminary Assessment of the Comparative Efficiencies of an SDI System Using Controlled or Natural Language for Retrieval. *Program*, 5:26-34; January 1971.

Pollitt, Steven. 1987. CANSEARCH: An Expert Systems Approach to Document Retrieval. *Information Processing & Management*, 23(2):119-138.

Saracevic, Tefko; Kantor, Paul; Chamis, Alice Y.; Trivison, Donna. 1987. *Experiments on the Cognitive Aspects of Information Seeking and Retrieving*. Springfield, VA: NTIS, 1987. (PB87-157699)

Saracevic, Tefko; Kantor, Paul. 1988. A Study of Information Seeking and Retrieving. II. Users, Questions, and Effectiveness. *Journal of the American Society for Information Science*, 39(3):177-196; May 1988.

Sewell, W.; Teitelbaum, S. 1986. Observations of End-User Online Searching Behavior Over Eleven Years. *Journal of the American Society for Information Science*, 37(4):234-245; July 1986.

Smith, Linda C. 1987. Artificial Intelligence and Information Retrieval. *Annual Review of Information Science and Technology*, 22:41-77.

Smith, Philip J.; Shute, Steven J.; Chignell, Mark H.; Krawczak, Deb. 1987. *The Role of the Human Factors Engineer in Designing the*

Interface to a Knowledge-Based System. Columbus, OH: Department of Industrial and Systems Engineering, The Ohio State University, 1987. (Report ISE-174)

Svenonius, Elaine. 1986. Unanswered Questions in the Design of Controlled Vocabularies. Journal of the American Society for Information Science, 37(5):331-340; September, 1986.

Vickery, Alina; Brooks, Helen M. 1987. PLEXUS--the Expert System for Referral. Information Processing & Management, 23(2):99-117.

Vickery, Alina; Brooks, Helen; Robinson, Bruce. 1987. A Reference and Referral System Using Expert System Techniques. Journal of Documentation, 43(1):1-23; March 1987.

APPENDIX A

FORMS FOR DATA COLLECTION

Moves Form
Selection Form
Reason Form

EXTRACTING KNOWLEDGE FOR INTERMEDIARY EXPERT SYSTEMS

MOVES FORM -- 1

Searcher ____________________ Search ________________________

Moves	# of times	ss#	Moves	# of times	ss#
Weight 1			Intersect1		
Weight 2			Narrow 1		
Weight 3			Narrow 2		
Weight 4			Intersect2		
Weight 5					
Negate					
Limit 1					
Limit 2					
Limit 3					
Limit 4					
Cut					

EXTRACTING KNOWLEDGE FOR INTERMEDIARY EXPERT SYSTEMS

MOVES FORM -- 2

Moves	# of times	ss#	Moves	# of times	ss#
Add 1			Expand 1		
Add 2			Expand 2		
Add 3			Expand 3		
Add 4			Expand 4		
Add 5			Exclude		
Include			Expand 5		
Cancel					
Refine			Probe 1		
			Probe 2		

Number of operationalist moves ________

Number of conceptualist moves ________

Total number of moves ____________

Operationalist _____%

Conceptualist ______%

EXTRACTING KNOWLEDGE FOR INTERMEDIARY EXPERT SYSTEMS

SELECTION FORM -- 1

Searcher ____________________

C & O	# of times	Searches
A.		
B.		
C.		
D.		
E1.		
E2.		
E3.		
E4.		
E5.		
E6.		
E7.		
E8.		
E9.		
E10.		
E11.		
E12.		
E13.		

EXTRACTING KNOWLEDGE FOR INTERMEDIARY SYSTEMS

SELECTION FORM -- 2

C & O	# of times	Searches
F.		
G.		
H.		
I.		
J.		
K*.		
L*.		
M*.		
K.		
L.		
M.		

EATRACTING KNOWLEDGE FOR INTERMEDIARY EXPERT SYSTEMS

REASON FORM --

Searcher	Search	Search key	Reason

APPENDIX B

OBSERVING AND INTERVIEWING ONLINE SEARCHERS

Nancy Phelps

To gather data for the National Science Foundation grant, "Extracting Knowledge for Intermediary Expert Systems: The Selection of Search Keys", the research team utilizes the case study method, a qualitative research technique. In this research, the method involves observing professional searchers performing five job-related searches, analyzing the searches according to models of moves and search key selection, and interviewing the searchers to clarify any misunderstandings. The three member research team consists of two graduate students and a professor, the principal researcher. The graduate students (hereafter referred to as the observers) perform the observation and interviewing tasks.

A unique element of the research design is the use of research methods commonly employed by social scientists--participant observation and personal interviewing. For years, qualitative methods have been accepted means of conducting research in the social sciences, since the social sciences are "human related" disciplines. Qualitative methods are new to research in the "hard" sciences or to research in technical matters. The data gathered as a result of this project will be used to advance the technical tasks of online searching and database construction, and the design of intermediary expert systems.

Use of qualitative methods is justified in this study because experienced online searchers are experts in online searching. To design effective intermediary expert systems and online databases it is necessary to understand how searchers perform their tasks, the logic behind the choice of strategies and search keys. To "get at" the searchers' knowledge, it is logical to study the searchers in real-life situations. Through observation, analysis and comparison of uncontrived searches, it is hoped that we will gain understanding of the process of online searching. The desire to understand the motivation of searchers' selections of search keys and strategies is related to the sociologist's desire to understand the motivations and feelings of the drug addict, for example. As the sociologist's goals logically dictate the use of qualitative methods, such as participant observation, so do the goals of this project.

In social science literature, there exists a body of writing on the problems encountered while using qualitative research methods. Though many of the problems encountered by this research team are similar, many have not been encountered and some new ones have been. It is my goal to elaborate the methods we use in our research, the problems that have been faced, and suggestions for coping with any difficulties. It is hoped that clarification of the qualitative methods used for online searching research will lead to further applications of the methods in other technical research.

STRUCTURE

We utilize qualitative methods of research. Although the theoretical roots of these methods are in the social sciences, we have adapted the methods to suit this unique situation. We utilize participant observation, like social science researchers. In traditional participant observation, the observer participates in a setting, such as an institution. The social scientist does not necessarily perform a job in the setting, but he/she actively participates in the situation and interacts with the organizational population. He/she is in the situation to observe behaviors, but observation is not his/her overt mission. He/she does not usually take notes in front of the observed individuals or tape record conversations. Recording and writing activities are usually done at the end of the observation day, after leaving the setting. The social scientist strives to "blend in", to not disrupt the normal interactions of the observation population.

There is a major difference in our methods of observation from the traditional ones. We are only interested in one individual performing one isolated task within an organization. The interactional dynamics of the organization are not important to us. Aside from common civilities, we are not interested in the other workers in the setting. We do not "blend in", and we do not disguise our interest in observing the online searching function. We openly tape record the search sessions.

A second difference in the methods is our use of interviews after completing observation and analyses of the searches. It is not unusual for social scientists to interview, but it is not standard to combine participant observation and interviewing. The participants in this study know they will be interviewed after completing the five searches. The interviews are to clarify any misunderstandings. If the searcher's choices are not clear, or if the comments made by a searcher are not consistent with the choices made, the issues are discussed during the follow-up interview. The interview session is formal and structured, deviating from the often loosely structured, informal, in-depth interviews of social scientists.

STEP 1: CONTACTING THE SEARCHER

In social science qualitative research, gaining access to an observable population is often difficult. Initial contact with qualified candidates is made by the principal researcher. The observers then contact the willing subjects. The observers only interact with searchers who have already agreed to participate, and who have some knowledge of the research purpose and design.

Eligible subjects are located through personal contacts, advertisements, and word-of-mouth. Personal contact seems to be the best method of securing participants, although all three methods work well. It has not been a problem to locate sufficient numbers of qualified, willing participants.

To date, there has only been one possible problem with this system. There have been a few cases in which the searchers' superiors were contacted in lieu of the actual participants. These searchers have been more nervous about their performances. They seem to question why they were chosen for participation by their superiors, and often feel that they are unqualified candidates. This problem has been handled indirectly, by ignoring the statements. In other situations, the negative statements have been contradicted by the observer. If it is not awkward to do so, it is best to ignore the insecure statements. The observer should not show any ability to judge the competency of the searchers. If the observer does respond, he/she should express their humble opinion. The observer should be sensitive to any expressed feelings of insecurity, and understand that this may be an ordeal for many participants. But to actively respond to expressions of doubt regarding a searcher's competencies may cause the observer to undermine his/her objective attitude. The observer should not only be perceived to be neutral, but should have a non-judgmental outlook. Responding to statements of doubt shows that the observer has the capacity to judge the searcher's performance, and that he/she can discern what are good and bad searches. That situation could change the dynamics of the searcher/observer relationship. The observer is in the situation to observe and possibly learn, not to compare one searcher to another.

The logical solution to the scenario is to make sure that each participant is personally contacted by the principal researcher before observation begins. Personal contact may eliminate the threatening feeling of being "picked" by one's superior, since the principal researcher asks the searcher to participate and expresses the team's need for qualified, professional searchers.

STEP 2: ESTABLISHING CONTACT BETWEEN SEARCHER AND OBSERVER

The observer contacts the searcher after initial contact has been made by the principal researcher. The observer only comes into contact with willing participants. The first contact between observer and searcher is made via the telephone. At this point, practical considerations determine the next step. If the searcher is located close to the University or the observer's home, arrangements are usually made to meet in person before the actual observation begins. If the searcher is located far from the observer, or does not have any extra time, the observer explains the procedure over the telephone. Whether meeting in person or not, it is important to be honest about the research goals but to remain vague regarding specific foci of the research. It is necessary to maintain vagueness because if the searcher learns before observation that the searches will be analyzed for strategical moves and selection of search terms, the searching behavior may be affected.

The explanation given to the searchers in this study is that the goal of the research is to observe five job-related searches (excluding known-item or classified searches) in order to understand the "art of online searching". I usually elaborate by explaining that the goal of the project is to make searching more accessible to begininning searchers and other interested individuals. I add that we want to discover what "goes on" in a searcher's head, what he/she knows that is not stated in books and manuals. The best way to accomplish the goals is to observe experienced online searchers performing actual searches. In this manner, I have adequately explained the research, the motivation for observation, the need for the searcher's participation, acknowledge the searcher's knowledge and expertise, but I have not revealed the specific foci of the study. I have been honest, but vague. It has been my experience that searchers are satisfied with this explanation and do not press for specifics. They are usually glad that formal research is recognizing their expertise and that the research design is practical. We are "asking" them what they do and think, instead of assuming a theoretical approach and telling them what they do and think. It has also been my experience that librarians are genuinely interested in helping each other, and in furthering knowledge in the field.

It is preferable, at this stage, if the observer can meet personally with the searcher. It is often difficult to establish a feeling of rapport over the telephone, and the feeling is essential to the success of the observation stage. If the individuals meet before beginning the formal sessions, there is less awkwardness when the observations begin. The introductions have already been made, the observer and searcher recognize each other, the observer knows the library lay-out and the location of the facility (he/she won't get lost on the road!) Any of the nervousness which may occur when two strangers meet is gone by the time the observation begins. You have met before, and the observer knows what to expect when entering the building.

Before beginning observation, it is important to explain to the searcher that he/she should speak freely and verbalize any thoughts that occur during the searching process. The searcher should never force comments or do anything that would feel unnatural. The observer should stress the unobtrusive nature of the observation.

As noted previously, it is extremely important to establish a feeling of rapport at this stage, whether meeting the searcher in-person or not. The observer should be friendly, polite and open. The observer must answer all questions without being overtly evasive (of course, if the searcher wants to know the specific focus of the study, it may be necessary to hedge). Be on time for the appointment, when meeting with the searcher. Dress neatly - you are entering a place of business. Do not make aggressive demands on the searcher. Accommodate the searcher's schedule. In this research, the observers are students and have class commitments. In this case, we inform the searchers of our general class schedule, without being specific - the searcher does not need to be burdened with our schedule. It is best to state that you are free Monday mornings, for example. The searcher should feel that his/her schedule is the important one, and that the observer will accommodate the searcher. Use good interpersonal skills. The observer should relate easily to many different types of people and be at ease with strangers and in unfamiliar surroundings.

It is not necessary for the observer and the searcher to become friends, and that would probably be detrimental to observation, but they should not actively dislike each other. We have not encountered this situation. Probably the only solution to an unfriendly pairing would be to cancel that searcher or use another observer. The searcher and the observer should relate well because they will be working closely through five searches. The situation may be unnatural enough for the searcher without the unnecessary tension of animosity. If personable observers are used in a study, this situation will probably never occur.

STAGE 3: OBSERVATION

Depending on the searcher's schedule and the frequency with which searches are requested in their organization, the actual observation period can last as little as one session or as much as several weeks. An observation session can last from one half hour to three or more hours, depending on the number of searches performed, their complexity, and the searcher's degree of preparation.

At this stage, the observer must change roles. The last (and probably also the first) contact with the searcher was a very personal, friendly introduction to the project and the observer, either in-person or over the phone. During the introductory session, the searcher may have expressed interest in the observer, what he/she is studying, career goals, etc. There may have been a friendly exchange covering many topics. Now the observer must try to assume a passive, unobtrusive role. The observer is present to observe, not to judge or comment. His/her full attention should be focused on the search. The observer should not suddenly become unfriendly. But he/she must withdraw and assume a very passive attitude. Establishing good feelings is no longer the aim of the observer. The observer must acknowledge by his/her attitude that the searcher and the searches are now of primary importance. The observer must become an active listener, but not an active participant. Active listening can be shown by nodding, saying "yes" when appropriate, and by not asking questions during the search.

If the observer has become too friendly with the searcher, he/she may find it difficult to "pull back" and become passive. The solution to the potential problem is to avoid becoming too friendly with the searcher during the introduction session. It may help to remember that the observer is entering the life of the searcher for a very limited time. The observer cannot become too emotionally involved in the relationship or his/her objectivity may be affected.

Observation is the most difficult phase of the project. The searcher may be very nervous, aware of the microphone, and may feel like he/she is being assessed. In response it is necessary for the observer to be neutral but supportive. To be supportive the observer should be aware of the possible emotions that the searcher is experiencing. It may also help if the observer is humble (probably an easy attitude for graduate students to assume) and willing to learn. If necessary the observer should explain that there are no right or wrong things to say while searching, and that the searcher may verbalize as much or as little as is comfortable.

Some well directed questions may calm the searcher, if he/she is extremely nervous, and push the observation session in the "right" direction. Ask what the search is about. This simple question will get the searcher talking about the search. Once the searcher begins talking, he/she may begin to feel less self-conscious.

It is best if the observer does not say too much, if at all possible. Silence may encourage the searcher to talk and explain what is happening in the search. The observer should not dominate the session by questioning the searcher. If the searcher does not talk while searching, the observer will probably still understand the logic of the search and the selection of search keys.

Each new observation session is usually less disruptive to the searcher. Over time, the searcher usually becomes accustomed to the observer and the microphone. Much of the nervousness of the early sessions will be gone by the end of the study.

Another possible problem encountered in the sessions is that the searcher makes typographical or other errors, possibly as a result of nervousness. Never point out the errors, no matter how helpful it would be to do so. If the observer points out errors, it places him/her in a judging role. Pointing out an error would place the observer in a position of being able to recognize errors and assess searches.

Occasionally the searcher may ask the observer if the searcher is "saying the right things". The searcher may express concern about what to say or what is expected from the session. The observer should be reassuring and noncommittal. It is good to reiterate the goals of the project as expressed previously. I have also found it useful to remind the searcher that I am a student and that the observation session is a learning process. I add that if it makes the searcher more comfortable he/she should consider me a pupil. The scenario seems to help dispel the nervousness and uncertainty of the searcher regarding what to say during the searches.

There may be a danger in observing people from one's own professional background. The observer may bring to the situation his/her biases or thoughts regarding correct and incorrect behavior and choices in the particular situation. In this research, there is no choice. The observers must understand online searching or they will not be able to analyze the searches. But it is not necessary for them to be expert searchers and it is best that they are not.

The problems of professionals observing their peers are mostly avoided through the use of graduate students. Though the student observers may bring certain searching biases to the sessions and may have contrary thoughts regarding searchers' choices, the observers know the limitations of their knowledge. Since they are students, the observers know they still have much to learn about searching. The naive attitude that the observer may need to assume while interacting with the searcher is easy and believable for a student to adopt. It is also plausible that a student will not comprehend all the intricacies of the searcher's logic, but it is not plausible for an experienced searcher or professor.

STEP 4: THE INTERVIEW

The interview is a very structured session. Unlike the observation stage, during the interview stage the observer actively participates, and directs the session. The observer must thoroughly prepare for the interview and know what questions are to be asked. It may even be helpful to write the questions, so that none are forgotten.

During the interview, it will become obvious that the observer is interested in the choices of descriptors and free text terms as search keys. To avoid awkwardness or an evasive attitude, the observer should inform the searcher of the questions' focus before beginning the interview. The explanation of the focus should be neutral. The observer may explain by saying, "I am interested in discovering when you decide to use descriptors to search and when you decide to use free text terms, so most of my questions will be on that issue." If the searcher asks for details, the observer may respond that neither choice of search terms is better, that the interest is simply when the choices are made, what conditions lead to a certain choice.

The observer should bring the search print-outs and searcher notes to the interview session. The searcher probably does not remember the specific searches in detail, and may need to see the print-out. The search analyses are not brought to the session. The searcher knows that his/her searches have been analyzed, but seeing the analyses may be threatening to the searcher. The interview may be difficult enough for some individuals without viewing the search analyses.

To avoid eliciting a defensive attitude from the searcher, the interview questions should be expressed with neutral wording. To achieve a neutral tone, it may be helpful for the observer to adopt a naive attitude. Phrases such as "I don't understand", or "I haven't seen this before, will you explain it to me", are neutral and express the attitude that the searcher is the individual with the knowledge. The searcher should feel that he/she is not being "grilled" or assessed regarding searching abilities.

During the interview sessions, I have heard several searchers express the belief that there is an ideal choice of search terms for a specific situation, or an ideal searching style, that the searcher is not utilizing. Many searchers believe that they are doing something in a search which is "against the rules", but I have never seen a searcher who follows the alleged "rules". Depending on the situation, it may be best to ignore these statements, thereby not giving them any credibility. It may be appropriate to subtly deny the statements, with phrases such as "you might be surprised at how others search". If the observer contradicts the statement, the response should be non-judgmental, and should not give the searcher the impression that the observer knows the "correct" way to search.

The observer's questions should be phrased using the pronoun "I" instead of "we", though the entire research team studies the search analyses and contributes questions. By using the singular pronoun, the observer relates the feeling that only a single person is viewing and analyzing the searches. If the plural pronoun is used, the searcher may feel that he/she is being closely scrutinized by a group of people. That may be a threatening feeling for many searchers.

The session should be controlled and subtly directed by the observer. The observer asks non-judgmental questions, allows the searcher to answer, and then asks for clarification if the searcher's response is not understood. The pattern is followed until all the questions have been asked and satisfactorily answered. It is best to not abruptly end the session. Allow the searcher to express any relevant thoughts. If it seems appropriate, at the end of the session the observer may ask the searcher for any general comments which were not covered by the questions. Since the questions follow a pattern, the searcher may have some comments regarding his/her own searching style. The searcher realizes that his/her searching style is being analyzed and understood through the observation and interview, and the searcher should feel that he/she is given a chance for fair representation.

At the end of the session, the observer should thank the searcher for his/her time. The searcher has done a favor by agreeing to participate in the study, and the favor should be acknowledged. The observer and searcher should leave the interview with continued feelings of friendliness. The searching community is small and any "bad blood" will probably circulate rapidly. We rely on the continued good will of the searcher, even when he/she is no longer a participant, as a walking advertisement for the project. If the searcher has found the research process pleasant and unobtrusive, he/she may encourage other eligible candidates to offer their services. Prospective candidates may question past participants, and it is important that the past participants convey a positive image of the process and the observers.

ATTRITION

To date, only one searcher has dropped out of the study. There is really nothing that can be done when the situation arises. If a searcher does not contact the observer for observation sessions, the observer should telephone a reminder. Before calling, the observer should allow about a week after the introduction session or the last search. Call in the middle of the work week. At the beginning of the week, a searcher may be too busy to think about the research. At the end of the week, the weekend is too close, and it may be difficult to schedule a suitable meeting time. The observer should not pester the searcher. The searcher has many tasks to perform, and his/her first priority is probably not the research. Contacting the searcher every seven to fourteen days is usually appropriate.

Even if the observer has done everything "correctly", the searcher may withdraw from the project with no warning or explanation. If this is happening, the observer may contact the searcher by phone and express interest in observing "x" more searches. If the searcher seems to have a negative attitude, do not push. The situation should be accepted and the conversation cordially ended. If the searcher responds positively to the call, but still does not contact the observer for observation, send a letter. Sending a letter is the least threatening gesture at this point. The searcher may not want to continue participating, but does not know how to withdraw gracefully. A letter can be ignored, and will provide the searcher an easy way to terminate. On the other hand, the searcher may be interested in continuing the project, but is pressed for time. A letter reminds the searcher of the project, but does not demand an immediate response.

The most important element to employ in conducting qualitative research is common sense. Any problems encountered during the research can usually be successfully handled by the researcher if he/she uses common sense and understands the feelings of the participants. Be polite, non-judgmental, empathetic, use common sense and your response will probably be effective.